THE UNIVERSITY *of* EDINBURGH

Intelligent Automatic Interpretation of Active Marine Sonar

John Charles Thomas Hallam

Ph. D. Thesis

University of Edinburgh

1984

Abstract.

This dissertation explores the problems raised by the design and construction of a real-time sonar interpreter operating in a three-dimensional marine context, and then focusses on two major research issues inherent in sonar interpretation: the treatment of observer and object motion, and the efficient exploitation of the specularity of acoustic reflection. The theoretical results derived in these areas have been tested where appropriate by computer simulation.

In the context of mobile marine robotics, the registration of sensory data obtained from differing viewpoints is of paramount importance. Small marine vehicles of the type considered here do not carry sophisticated navigational equipment, and cannot be held stationary in the water for any length of time.

The viewpoint registration problem is defined and analysed in terms of the new problem of motion resolution: the task of resolving the apparent motion of objects into that part due to the movement of the observer and that due to the objects' proper motion. Two solutions to this underconstrained problem are presented. The first presupposes that the observer orientation is known *a priori* so that only the translational observer motion must be determined. It is applicable to two and three-dimensional situations. The second solution determines both the translational and the rotational motion of the observer, but is restricted to a two-dimensional situation. Both solutions are based on target tracking techniques, and have been extensively tested in two dimensions by computer simulation. The necessary extensions to deal with full three-dimensional motion are also discussed.

The second major research issue addressed in this thesis is the efficient use of specularity. Specular echoes have a high intrinsic information content because of the alignment conditions necessary for their generation. In the marine acoustic context they provide a significant proportion of the information available from an acoustic ranger. I suggest a new method that uses directly the information present in specular reflections and the history of the vehicle motion to classify the specular echo sources and infer the local structure

of the objects bearing them. The method builds on the output of a motion resolution system. Six distinct types of specular echo source are described and three properties useful for their discrimination are discussed. A suitable inference system for the analysis and classification of specular echo sources is also proposed.

Table of Contents.

List of Sections.

List of Symbols and Abbreviations.

$\cdot$	Vector dot product operator.
$\times$	Vector cross-product operator.
A^T	Transpose of the matrix A.
A^{-1}	Inverse of A.
detA, **det**(A)	Determinant of A.
var[ ]	Scalar variance.
E[ ]	Expected value.
cov[a b]	Covariance of two scalars.
cov[$\underline{a}$]	Covariance matrix of a vector.
$\underline{v}_O$	Observer proper linear velocity.
$\underline{\omega}_O$, ω_O, $\underline{\omega}$, ω	Observer angular velocity.
${}_P\underline{v}_O$	Velocity of P with respect to O.
θ	Observer polar axis offset angle.
$\delta\omega$	Differential or error observer angular velocity.
$\delta\theta$	Observer polar offset error.
$\underline{rm}(\underline{r})$	Reflected observer motion at position $\underline{r}$.
$\underline{x}_i$	Target or observer relative state vector.
$\underline{\xi}_i$	Target or observer absolute state vector.
ε	Affine analysis expansion factor.
λ_{max}, λ_{min}	Affine fit error eigenvalues.
$\delta\underline{v}$, δv	Linear velocity error.
CR	Constrained relationship metaphor for segmentation.
EG	Explanation generator metaphor for segmentation.
IAF	Information Averaging Filter.
OAS	Observer absolute state.
PEV	Position error variance.
TAS	Target absolute state vector.
TRS	Target relative state vector.
rd	radian.

SI units are used throughout this dissertation, and vectors, unless otherwise noted, are assumed to be single column matrices.

Chapter 1. Introduction.

Most contemporary work on intelligent machine vision systems is concerned with the interpretation and understanding of images collected by optical sensory equipment. In air, optical sensing allows long range high resolution exploration of an environment to be carried out. In the marine environment the preferred sensory mechanism is acoustic. There is little ambient light in the ocean and optical visibility is poor. The physical properties of the medium favour acoustic sensing and the ability to determine range directly from time of flight measurements provides a further attraction for sonar (SOund NAvigation and Ranging) systems.

This dissertation describes an approach to the problem of interpreting data collected by acoustic ranging systems where these are used as the (or a) principal sensor for a robotic vehicle. I shall be concerned with active sonar systems -- sensors which repeatedly insonify the environment and then detect the echo returns from the structures in the neighbourhood of the transducer -- rather than passive sonar sensing, which relies on noise emitted by the structures to be discerned and has been the focus of the majority of military sonar investigations.

The sonar interpretation problem I shall be discussing was motivated by the work currently being done (and work envisaged for the future) by the Underwater Technology Group in the Department of Electrical and Electronic Engineering at Heriot-Watt University, Edinburgh, on their "ANGUS" submersible vehicle. A brief description of the ANGUS vehicle is given in the next section, as background for the treatment of sonar interpretation. The scope and content of this thesis are described in section 1.3, and section 1.4 is a short guide to the organisation of the thesis.

1.1 Background.

The ANGUS vehicle (Dunbar and Holmes, 1978) is a prototype unmanned tethered submersible designed for marine inspection and survey applications at a typical operating depth of 300 to 400 metres. The vehicle carries some navigational equipment (detailed below) and can carry television cameras and various sonar equipments. It is also capable of carrying a manipulator or a small free-roving satellite vehicle.

The existing vehicle is controlled by a shipboard operator via a minicomputer to which it is connected by the tethering umbilical cable, which carries power and control signals from ship to vehicle and television and sonar video from ANGUS to the ship. The navigational equipment carried by the existing vehicle (ANGUS 002) comprises:

- a flux-gate compass and echo depth sounder that provide height and heading information for the vehicle guidance and control system;

- a SIMRAD HPR-205 short-baseline active transponder navigational sonar, that computes the vehicle position by triangulation with respect to two active transponder beacons fixed near the seabed at the ends of a baseline (this information is also used by the guidance and control system);

- and a sector scanning sonar that provides a display for the vehicle operator and sense data for automatic control functions.

Although ANGUS is currently attached to its mother ship by cable it is intended as a long term goal to replace that physical link with an acoustic data link. The satellite vehicle which ANGUS will carry is a mid-term goal in the project, and an experimental version is currently under construction. It will communicate with ANGUS by

acoustic link and will be used for work in cramped or cluttered environments where the umbilical cable could easily become snagged. As a very long term goal it is intended to provide intelligent software that would enable ANGUS to perform sophisticated tasks, for example object recovery, with little or no supervision from an operator.

There are four project goals associated with the use of sonar data in the ANGUS project.

1). The sector scanning sonar provides information about the presence of obstacles in the vehicle's neighbourhood and indicates their approximate size and extent. In the short term it is intended to extract this information and make it available to the vehicle guidance system so that automatic collision avoidance and targeting (swimming to a given goal position) can be implemented.

2). Sonar is currently presented conventionally to the operator as a radar-like plan position display of echo amplitude against range and bearing plotted in polar coordinates. This type of display places a significant load on an experienced operator and it is desirable to display sonar in a form more easily interpreted by both novice and experienced operators.

3). In the long term, when the umbilical cable carrying sonar and television data is replaced with an acoustic link, there will be a serious bandwidth problem. The acoustic link has a bandwidth of about 10kbit/second, so sonar and television video data alike must be substantially compressed if the link is to carry them.

4). In the very long term ANGUS will be expected to perform a variety of intelligent activities using interpreted sensory data, for example, to recover objects or plan trajectories for avoiding obstacles, probably with very little supervision from an operator.

In summary, the open-ended nature of the ANGUS project allows interest in a variety of intelligent tasks which could be automated, depending not only on sonar data but additionally on data from other senses, for example tactile and television information. Such tasks will be implemented as and when their contribution to the project becomes significant. In the short term, intelligent use of sensory data could be made to present a composite synthetic display to the operator, showing control information, television images and interpreted sonar data superposed to create an easily understandable whole. In the long term the interpretation of sense data would help in data compression and would drive other intelligent activities such as object recognition or trajectory planning.

1.2 The Scope and Content of this Thesis.

The fundamental intelligent activity underlying both the short and the long term goals of the ANGUS project is the primary interpretation of acoustic sensory data. This dissertation explores the problems raised by the design and construction of a real-time sonar interpreter operating in a three-dimensional marine context, and then focusses on two major research issues inherent in sonar interpretation: the treatment of observer and object motion, and the efficient exploitation of the specularity of acoustic reflection. The theoretical results derived in these areas have been tested where appropriate by computer simulation.

The discussion of the sonar interpretation problem is grounded in the context of marine mobile robotics. It is assumed that the sonar data to be interpreted is collected by a submersible vehicle of the same generic type as ANGUS, and that the interpretation of the data is intended to provide a basis for intelligent control or display. However, much of the discussion of observer and object motion is applicable in the more general context of mobile robotics.

1.2.1 The Sonar Interpretation Problem.

Sonar interpretation consists in using data collected by an acoustic ranging sensor to construct a computational description of the environment. If the resulting description is to be useful for intelligent display, or for task planning and control applications, it must include a detailed three-dimensional geometric description of the shape of the local seabed and of significant structures (objects) near the sensor. It must also describe the positions and velocities of the objects and of the vehicle carrying the sonar transducer. Since the vehicle will move continually, the description should be recorded in a viewpoint independent way.

This purely descriptive function differs from scene understanding programs such as Brooks' ACRONYM system (Brooks, 1981) or Fisher's IMAGINE (Fisher, 1983), both of which attempt to recognise instances of generic (ACRONYM) or particular (IMAGINE) objects in the scenes presented to them. The recognition task tackled by these programs involves the invocation and incremental instantiation of *a priori* models of the environmental structures that the program is attempting to recognise, using data from the sensors. Sonar interpretation, as defined here, implies no recognition of *a priori* objects but rather the construction of *a posteriori* models of the environmental structures actually seen.

The sonar interpretation process is closely related to the problem of constructing a suitable intermediate representation from an optical image. Such a representation typically forms the basic input for "high-level" visual processing such as recognition systems, and is thus analogous to the output of sonar interpretation.

The problem of constructing an intermediate visual representation has been discussed, for example, by Marr (1982), who proposed an observer-centred (propriocentric) description comprising relative depth, surface orientation and surface discontinuity for the environmental surfaces. The description, which he called a $2\frac{1}{2}$D sketch, includes the computational mechanisms necessary for maintaining the various components of the description in a mutually consistent state

as the view being described changes.

The sonar interpreter output differs in a fundamental way from the $2\frac{1}{2}$D sketch of Marr, however, since the latter is recorded in a propriocentric way while the former is a viewpoint independent description.

There are cogent reasons for requiring the description generated by sonar interpretation to be viewpoint independent. First, the surface properties of objects (such as shape) and their proper motions are properties of the objects themselves and are independent of the observer (although clearly the observer's perception of these properties is propriospecific). Thus as far as possible the description of these properties should be object-centred (exterocentric) rather than observer-centred (propriocentric). Second, the observer viewpoint is continually changing as the robotic vehicle carrying the sensor moves about. In the marine context it is impossible to hold a vehicle stationary, so the registration of different propriocentric descriptions or the transformation of a propriocentric description must happen continuously and in real-time. Third, most of the users of the results of sonar interpretation are tasks that are naturally specified in exterospecific terms. It is inefficient to require several users of the results of sensory interpretation each to convert the description from a propriocentric to an exterocentric form before using it.

An alternative perspective on sensory interpretation has recently been suggested, in the context of visual interpretation, by Witkin and Tenenbaum (1983). They propose that the organisation of sensory data into perceptual units is not merely descriptive -- it is principally explanatory. Thus, for example, the perception of certain combinations of two-dimensional motions among the components of an image as projected movement of a rigid body in three dimensions is not primarily a concise description of the image motion; rather, it is an assertion by the sensory interpreter that the two-dimensional motions were indeed caused in that way. The sensory interpreter postulates an appropriate rigid body to explain its sensory input.

In the case of rigidity, this perspective complements the approach taken by Ullman (1979), who postulated rigidity as an axiom for the recovery of structure from motion. He required the environment to have the property that any motion of a set of image elements that could be interpreted uniquely as a three-dimensional motion of a rigid body was in fact caused in that way.

The picture of sensory interpreters as generators of explanations is an appealing one for the sonar context because it describes well the task of sonar interpretation -- the elucidation of a model of the environment from the sensory input. This a posteriori model can legitimately be regarded as an explanation of what has been observed, as well as being a description of the environment.

The explanation generator paradigm also highlights the common ground between the function of a sonar interpretation system and that of a recognition-based scene understanding system such as ACRONYM, in that both are attempting to explain their sensory input in a model-orientated way, the former by constructing a posteriori models of "low-level" entities such as surfaces, edges, corners, and so on, while the latter instantiates a priori models of complete objects.

In summary, sonar interpretation is the process of elucidating a geometric model of environmental structure from sonar data. I shall therefore use the term sonar interpretation to mean:

> the use of acoustic range data, gathered by a mobile marine submersible vehicle, to construct detailed exterocentric three-dimensional computational models of the shape of the seabed and objects in the vehicle's environment and to deduce the positions and velocities of those objects and the observer with respect to a fixed viewpoint independent frame of reference.

A sonar interpreter performs this task.

1.2.2 Observer and Object Motion.

Interpreting sensory data in the marine context requires a new paradigm of motion. Land-based robotic navigation systems are able to assume that the vehicle is the agent of its own motion -- that it will remain stationary unless a deliberate movement is made. In the marine context this active motion paradigm is inadequate because motion has a significant passive component. Marine currents perturb the vehicle position and movement is continually impressed on the robot by its environment. For robots which, like ANGUS, do not carry sophisticated navigational equipment and cannot obtain precise position information, this continual motion creates a serious problem for a sensory interpreter that must assimilate data obtained at different times from different viewpoints.

I describe a novel paradigm suitable for the sonar interpretation problem, incorporating both active and passive motion by recovering both the observer and object motions from the relative motion of objects. The determination of the observer motion is cast in terms of a novel problem, the task of resolving the apparent motion of an object with respect to the observer into that component which arises from the motion of the object and that component caused by the motion of the observer. This problem will be termed the motion resolution problem.

A solution to the motion resolution problem is presented, based on the type of information available from an acoustic ranging sensor. The input to the solution consists of a time series, possibly irregular, of noisy measurements of the relative positions of a variable number of feature points in the environment. These points are assumed to be on environmental objects. No use is made of the structure of those objects and the solution does not require knowledge of the association between points and objects. It does, however, require knowledge of the correspondence between measurements and feature points.

The solution presented employs conventional target tracking techniques to obtain estimates of the current position and velocity of the feature points. These estimates form the input to a new algorithm that resolves the estimated motion into the component caused by the observer motion and that attributable to the motion of the feature point.

Two versions of the motion resolving algorithm are given. The first is formally defined as a linear recursive statistical filter, and is applicable to situations where the motion of the observer is translational, or where the orientation of the observer can be determined independently of the motion resolution processing. The algorithm is described in a two-dimensional situation but is applicable to an arbitrary number of dimensions.

The second version of the algorithm is an extension of the first to deal with an observer capable of full two-dimensional motion, i.e. having one degree of rotational freedom and two degrees of translational freedom. It is only applicable to this situation as it stands, but can be extended to handle full three-dimensional motion of the observer. The necessary extensions are also described.

1.2.3 Exploiting the Specularity of Acoustic Reflection.

The second major research issue addressed in this thesis is the efficient use of specularity. Acoustic reflection in water generates a larger proportion of specular reflections (highlights) than optical reflection because of the physical properties of the medium and the objects being observed. Acoustic ranging systems make use of these specular reflections as well as the more familiar diffuse ones. However, because of the greater proportion of specular reflections much of the environment may be invisible to the transducer at any given time.

In optical sensing, specular reflections may be avoided by using matt object surfaces or carefully positioned illumination. The specularity inherent in acoustic ranging may also be circumvented, by using special acoustic imaging sensors. Acoustic imaging systems

provide an optical style image, sometimes including range information, by means of acoustic lenses, aperture synthesis, or holographic techniques (Sutton 1979). They are appropriate for detailed inspection of objects at close range (within the Fresnel zone of the transducer) but require large processing resources.

Specular echoes, however, have a high intrinsic information content. In common with diffuse echoes they permit the sonar equipment to determine the relative position of the echo source, but they also constrain the local structure of the object responsible for the echo (because of the alignment of the observer and object necessary for a specular reflection to occur). Rather than generating acoustic images and applying conventional image processing or scene understanding techniques to deduce the geometric and topology of the object surfaces, I suggest a new method that uses directly the information present in specular reflections and the history of the vehicle motion to classify the specular echo sources and infer the local structure of the objects bearing them. The method builds on the output of a motion resolution system.

I present an enumeration of the properties of six distinct types of specular echo source. The sources are classified in terms of three properties that can be deduced by reasoning about the relationship of the apparent trajectory of the echo source to the trajectory of the observer. These are the topological structure of the set of positions from which the source is known to be visible, the structure of the set of points known to lie on the source, and the proper motion of echoes known to be generated by the source.

Methods of obtaining a description of the set of viewpoints from which a source is visible are discussed. A statistical technique for inferring the second property is given and its behaviour is tested in certain simple cases using a computer simulation. The third property can be tested directly using the output of the motion resolution system. A suitable inference system for the analysis and classification of specular echo sources is also proposed.

1.2.4 The Use of Computer Simulation.

The motion resolution algorithms referred to above were tested extensively by computer simulation in order to assess their performance. There were several reasons why a simulation was chosen for the evaluation of the method rather than a test using real sonar data.

First, real sonar data, (collected, for example, by the ANGUS vehicle) contains full three-dimensional observer motions and would therefore require extensive preprocessing if the motion resolution problem were not to be solved for three-dimensional motion in one step. Using simulated data allows the class of observer motions to be restricted and precisely controlled.

The use of real sonar also requires a correspondence problem to be solved. The echoes detected by a sonar device are anonymous and irregular. It is usual for a given target to fade intermittently, so there is no guarantee that any particular sonar scan will contain an echo from that given target, and this further complicates the correspondence problem. In fact, this correspondence problem is best solved in conjunction with the estimation of the relative velocity of targets by a target tracking system, but it represents an additional complication for real data testing.

The evaluation of the performance of complex statistical algorithms such as the motion resolution algorithm described in this thesis is normally carried out using a Monte Carlo sampling technique. The algorithm is run on a set of randomly generated test cases and its performance on the sample is taken to be indicative of its performance in general. The statistics reported in Chapters four and five were obtained from test runs using typically 350 simulated observations of the targets, so to provide a single Monte Carlo test set equivalent to the ones used there would require about 7000 frames of sonar data.

The simulation, unlike real data, is precisely controllable, and the target and observer motion parameters can be chosen at will. This makes it possible to test particular parts of the algorithm without involving others (such as the angular velocity extraction algorithm in Chapter five, for example). The simulator is also able to supply the true position and velocity information associated with any target or the observer and this data provides the basis for the error analysis of the algorithm.

The simulation used for the evaluation of the algorithms was designed to be as simple as possible while retaining the salient characteristics of sonar derived data. The simulation was restricted to two dimensions so that a simplified version of the motion resolution problem could be considered. Objects, of which the observer is a special case, were treated as points having two degrees of freedom corresponding to their position coordinates in a world coordinate system (except that, in the tests described in Chapter five, the observer was permitted a degree of rotational freedom also). At intervals specified by the sonar model the positions of certain objects relative to the observer were made available to the algorithm under test; these measurements could be corrupted by the addition of controlled noise if required. To model the intermittent visibility that is characteristic of sonar a probability parameter was defined for each object except the observer. This was used to determine whether the object was visible during any particular sonar scan.

Objects were allowed to move in a purposive manner, i.e. they were assumed to follow a predetermined trajectory with an optional random perturbation of their position. The basic trajectories were described by a linear state transition model and included piecewise linear constant velocity tracks. The predetermined track was designed to model both deliberate motion of the observer or target and the motion of objects caused by large scale water currents, while the random component of the trajectory modelled the perturbation of the object motion by turbulence and small scale water currents.

1.3 Reader's Guide.

The material presented in this thesis has been arranged in such a way that it should be easy to read straight through. This section describes the overall structure of the thesis, and is intended both to give an overview of the way the chapters interrelate and to enable selective readers to find quickly the chapters relating to their particular interests.

Chapter two concentrates on the process of sonar interpretation, enumerating the constraints that arise from the interaction of a sonar interpreter with its physical environment. It suggests a suitable modularisation for a sonar interpreter and briefly explores the function of each module. The motivation is then given for the two specific modules considered in the remainder of the thesis.

Chapters three to six constitute the discussion of Viewpoint Registration, which is the problem of constructing a viewpoint independent frame of reference and of maintaining a transformation from the current viewpoint dependent reference frame associated with the sonar sensor to the exterocentric reference frame.

Chapter three introduces the motion resolution problem, of crucial importance to Viewpoint Registration. This is the task of resolving the apparent motions of objects into the component caused by their proper motion and the component caused by the motion of the observer. The chapter concludes with a critical review of work relevant to the problem in the areas of robot and terrain-aided navigation, optical flow, and radar or sonar target tracking systems.

In Chapter four a solution to the motion resolution problem is described informally and defined formally in terms of recursive linear estimators such as the Kalman filter. This solution is constructed around a target tracking system and it applies to robotic vehicles with two translational degrees of freedom of motion, or whose orientation in the plane can be determined independently of the motion resolution processing. The formal definition assumes a moderate degree of familiarity with the theory of statistical

estimation. Readers wishing to grasp the operation of the solution without necessarily understanding it deeply are advised to skip the indicated sections.

Chapter five extends the motion resolving system developed in Chapter four to situations where the observer can execute full two-dimensional motions. It ends with a summary of the conclusions drawn from the testing of the full two-dimensional motion resolution system and gives several suggestions for future, related, research.

Chapter six explores in greater depth two of the suggestions given at the end of Chapter five. It discusses the extensions necessary to the motion resolution system of Chapter five for handling full three-dimensional motion, and describes mechanisms for incorporating direct observer motion information (obtained, for example, from the vehicle's control system) into the Viewpoint Registration processing.

Chapter seven deals with the segmentation problem associated with Viewpoint Registration in terms of a critical review of state-of-the-art techniques for data association in radar and sonar target tracking systems. Some suggestions for the segmentation processing implicit in Viewpoint Registration are also made.

In Chapter eight, the consequences of the specularity of acoustic reflection are explored. A classification of the different types of echo source in terms of their visibility and relative motion properties is given and some techniques are suggested for obtaining these properties from histories of echo movements.

Chapter 2. Designing a Sonar Interpreter.

Sonar interpretation has been defined as a computational process that uses acoustic range data to construct and maintain a detailed geometric model of the environment. The sonar interpretation problem has further been set in the context of mobile marine robotics.

This chapter explores the constraints on the sonar interpretation process entailed by its interaction with the physical environment it is attempting to describe. A modular decomposition of the problem is proposed and the research issues associated with each module are described briefly. The chapter concludes by presenting the motivation for focussing on the Viewpoint Registration and Specular Event Analysis modules, which constitute the subject of the remainder of the thesis.

2.1 Constraints on Sonar Interpretation.

The constraints that influence the design of the interpreter can be grouped (albeit somewhat fuzzily) into two classes: constraints of implementation and constraints inherent in the computational process implemented by the interpreter system. The former result from the attempt to implement sonar interpretation on a given set of hardware, or in a given computational environment, or indeed from the intention to implement the process at all, and predominantly affect how the implementation must be done.

An obvious implementation constraint is response time: if a system is to provide useful feedback to an operator, for example, a response time of more than a couple of seconds is unacceptable. The data rate of the sonar driving the system is another constraint of the implementation which, in relation to the speed and power of the hardware, determines what processing can be done in a given time.

A less obvious implementation constraint is, for example, that the viewpoint independent output of sonar interpretation should be recorded as a concurrently accessible database. This constraint is

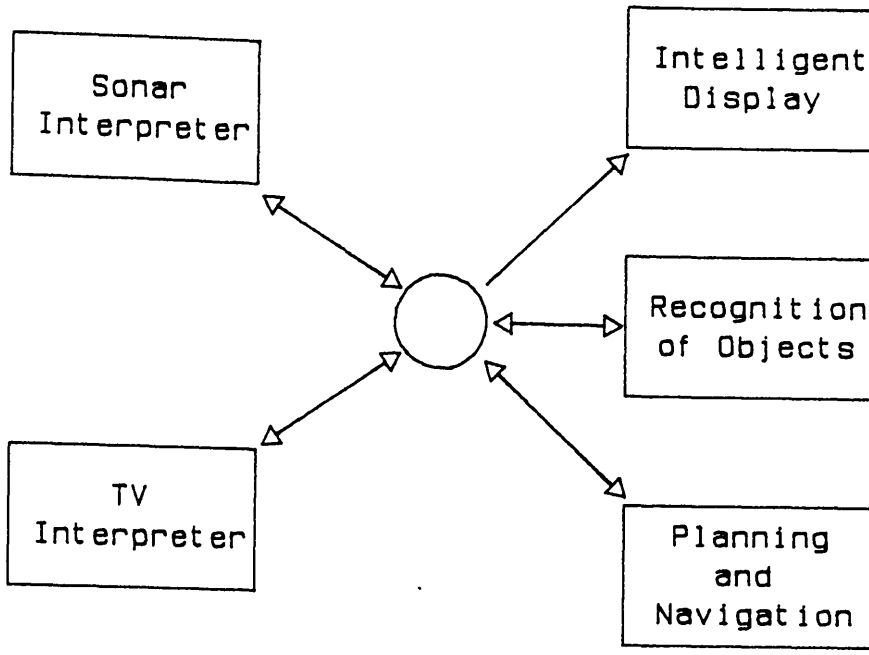


Figure 2.1.

suggested by the potential interactions between intelligent sensory interpretation systems and the users of interpreted sensory data. An example, illustrated in Figure 2.1, is automatic object retrieval. The sensory data is provided by cooperating sensory interpreters, in this case say for sonar and television, and is used concurrently by a recognition task and a navigation or search-planning task; an intelligent display for the vehicle supervisor is also shown in the figure.

Implementation constraints such as these affect the choice of algorithms used to express the various processes internal to the interpreter, but they do not provide a fruitful insight into the structure of the sonar interpretation process and will not be considered further here.

Inherent constraints on the sonar interpretation process arise from two sources: the mobile marine context in which the problem has been set, and the physical process by which the sensory data is acquired for interpretation. These areas are considered below, and the various constraints they generate are identified.

2.1.1 The Contextual Constraints.

There are three constraints arising from the context in which the sonar interpretation problem has been set. All are consequences of the behaviour of small marine submersibles.

In the ocean it is impossible to keep a small vehicle stationary with respect to the world so that sensory data can be gathered from a stable viewpoint; furthermore, inexpensive robots like the ANGUS vehicle do not carry the sophisticated navigational equipment necessary to give precise information concerning the vehicle's position and movement. This creates a serious data registration problem since the sonar data to be interpreted is gathered from a continually changing viewpoint, and must be transformed into a viewpoint independent description or explanation of the environment using the imprecise knowledge of the position and motion of the viewpoint supplied by the vehicle's navigational or control system. The conundrum is expressed by a pair of constraints.

- The input to the sonar interpretation process comprises viewpoint dependent (observer-relative) positions of objects in the environment.

- The sonar sensor viewpoint changes continually and unavoidably as the vehicle moves and only imprecise information is available describing its motion.

The latter constraint entails another which, although strictly speaking an implementation constraint, influences the modularisation of the interpretation process and constrains the realisations of the modules. The constraint is this.

- Sonar interpretation is essentially a real-time process.

This constraint becomes important, for instance, in the discussion of the Viewpoint Registration module described in section 2.2.2.

2.1.2 The Physical Constraints.

The second class of constraints arise out of the nature of the physical process used to gather sense data for interpretation. In the marine environment the most suitable process to use for sensing is acoustic reflection. There are two main reasons for this.

First, the physical properties of the medium favour acoustic sensing. The relatively high speed of sound (about 1500 metres per second), the reasonable attenuation properties of the ocean, and the ease with which water can be insonified, make possible the construction of long range sensors with considerable resolution. A typical commercial acoustic ranging system might have a maximum range of about 500 metres, a beam width of two degrees of arc, and a range resolution of perhaps twenty centimetres.

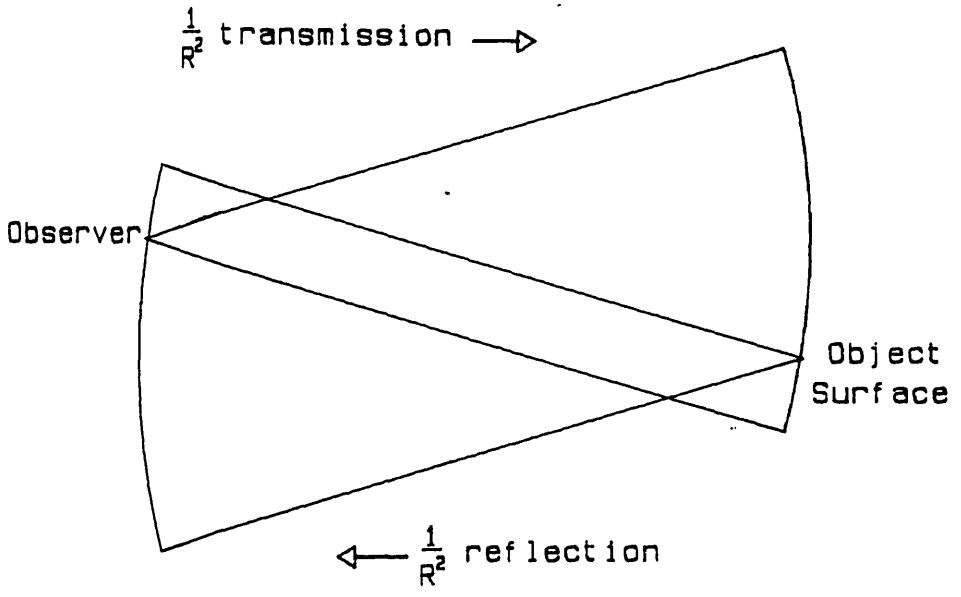


Figure 2.2. Inverse Fourth Power Attenuation for Optical Sensors.

Second, optical sensors are disadvantaged by the lack of ambient light, which necessitates an active optical system with the inverse fourth power law that that entails (see Figure 2.2), and by the low visibility (often as little as 20 cm in the operating environment of ANGUS) caused by dirt in the water.

Acoustic sensing has a number of unusual properties in relation to the more familiar optical imaging process, and these give rise to constraints on the sonar interpretation process in the same way that the physics of optical imaging constrains image interpretation.

The most obvious difference between a sonar sensor and an optical sensor such as a television camera is that the sonar sensor provides a direct measurement of the relative range and bearings of any echo sources it detects. Its output is thus intrinsically three-dimensional. In this respect it is similar to laser rangefinders or triangulation depth sensors, for example.

However, a second difference is that acoustic reflection is highly specular relative to optical reflection. The interfaces between the ocean and the inanimate objects in it are acoustically shiny while the relatively long wavelength of sound (in comparison to optical wavelengths) used by sonars allows the surfaces of objects to appear, acoustically, relatively smooth. Together these effects make strong specular reflections commonplace (although diffuse reflection is also common). Concavities in objects, such as concave corners and edges, are especially strong reflectors.

This tendency to specularity has two implications which make sonar images very different from optical images. Consider the situation in

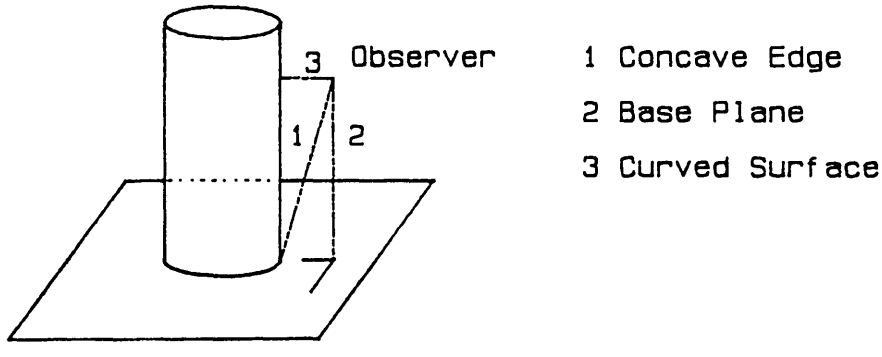


Figure 2.3. Specular Image of a Cylinder and Base Plane.

Figure 2.3; the observer at O is scanning an acoustically smooth, shiny, cylinder. The sonar image comprises a strong specular reflection from the concave edge where the cylinder meets the base plane; a somewhat weaker reflection from the base plane below the observer;

and a weaker reflection from the cylindrical surface (the reflected intensity falls with increasing surface curvature). There will also be weak reflections where sound waves pass the cylindrical surface (Freedman, 1962). Most of the object and base surfaces are invisible.

In practice, surfaces are not perfectly smooth acoustically and there will be glints from the surface irregularities. Thus the concave edge would probably generate several glints from uneven places along its length, and extra points on the base and cylinder would be visible. However, much of the object surface still remains unseen.

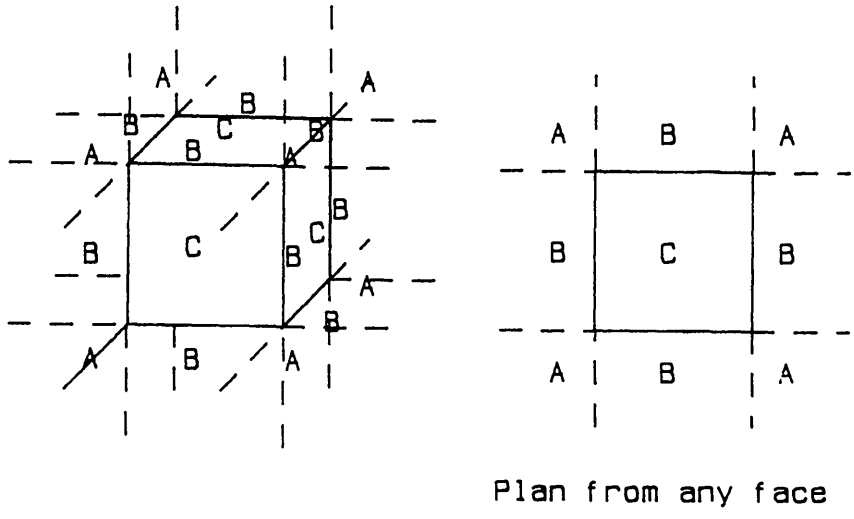


Figure 2.4. Viewpoint Sets Associated with a Convex Corner.

The second effect of specularity is that the image changes catastrophically as the observer moves relative to the object. Figure 2.4 illustrates this for a convex corner. If the surfaces are acoustically smooth and shiny, the set of possible viewpoints is divided into regions in which different types of images obtain, with catastrophic image changes at the boundaries. In the type A region the weak reflection from the convex corner is all that is visible; in type B regions, a weak reflection from an edge can also be seen; in type C regions the strong specular reflection from a face is present as well. Again, in practice, the situation is improved by the roughness of the surfaces. However, the weak reflections may not be resolvable from the background noise and there can be extensive regions of space in which the corner is acoustically invisible.

The presence of a number of good reflectors of sound (and the boundary at the sea surface is one such excellent reflector) and the refraction of sound as it passes through regions of varying water pressure, salinity or temperature, make multiple transmission paths from the sensor to its target a fairly common occurrence. This causes multipath distortion and false detections. In part it also accounts for the frequent fading of sonar -- where a potentially visible object is randomly acoustically invisible -- which is probably due to the constructive and destructive interference of echoes from multiple paths and different parts of the object.

A final property of acoustic sensing relevant to sonar interpretation is the phenomenon of "range shadows". Although the greater part of an object's surface may be acoustically invisible from any given viewpoint, the object still casts a shadow, which may be detected by the variation with range of the ambient noise received by the sonar sensor.

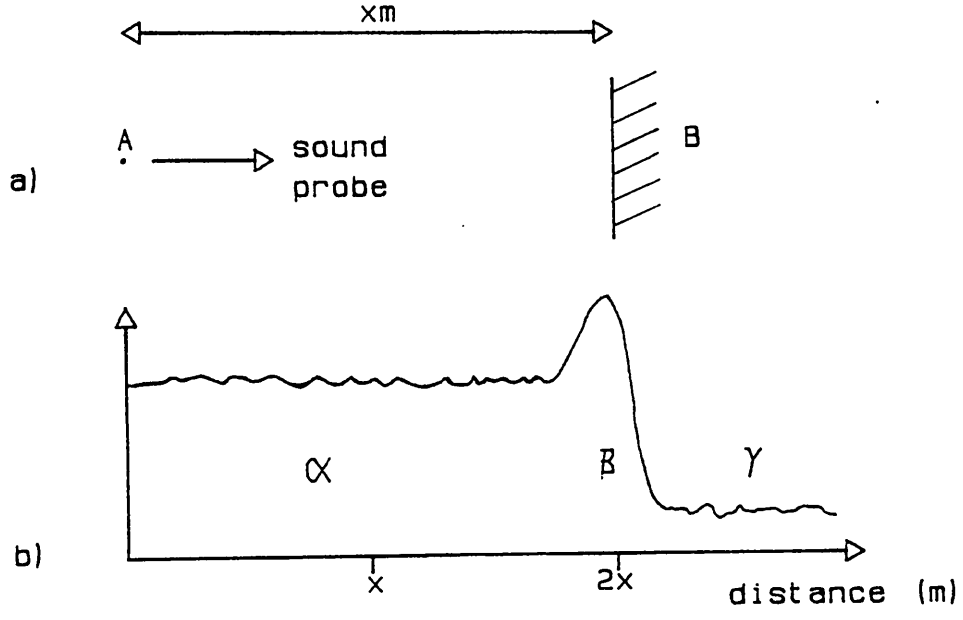


Figure 2.5. The Generation of Range Shadows.

Consider the situation in Figure 2.5a, where the sonar at A has insonified its environment and is attempting to detect echoes from the direction of the barrier B. As the sound wave passes through the water between A and B, some acoustic energy is scattered back, causing an increased ambient noise, shown as region α in Figure 2.5b.

When the sound wave reaches the barrier, it is reflected back, causing the large pulse β in Figure 2.5b. However, once the wave has passed the barrier (neglecting diffraction effects) the water is effectively no longer insonified and the noise level falls, as in region γ. (Note that I have "unpacked" the folded reflection path in the diagram for ease of explanation.)

This effect shows up on a sonar display as a shadow behind the bright echo caused by the barrier. Such range shadows are of great importance in human interpretation of sonar because they convey an impression of the size and disposition of objects in the environment. The shadows are not sharp, as optical shadows tend to be, because the relatively large wavelengths used encourage diffraction effects, and unlike optical shadows (for the case when the source of illumination is coincident with the observer) the obscuring surface may be acoustically invisible because of specularity.

These properties of acoustic sensing entail a number of constraints on the design of the sonar interpreter. Multipath distortion, fading, and the positioning of space due to specularity imply that

- a given object point is visible only intermittently, at random,

while the strength and frequency of specular reflections suggest that

- the interpreter must make use of the information present in both specular and diffuse echoes

and

- for much of the time large sections of the object surface will be effectively invisible.

In this situation range shadows contain useful size and shape cues, so

- the interpreter must be able to detect and analyse range shadows.

Finally, the environment is a noisy one -- air bubbles and fish contribute to the spurious detections ("clutter") presented to the interpreter, and the signal to noise ratio of the sonar equipment is limited by cavitation at the transducer. Thus,

- The interpreter must degrade gracefully as the information content of its input falls, and it must be robust enough to cope with noisy data.

2.2 The Components and Organisation of a Sonar Interpreter.

The various constraints identified above can be used to decompose the sonar interpretation process into subtasks. A number of subcomponents are evident: the interface to the sonar equipment driving the interpreter; Viewpoint Registration; Echo and Shadow analysis; Modelling and Database maintenance. These subtasks are partially ordered by their input requirements: the Sonar Device Interface converts the raw data from a sonar equipment into an internal representation of echo detections (events) and shadow regions that is used by the rest of the interpreter; Viewpoint Registration is a necessary precondition for the Analysis and Modelling; the two aspects of Analysis, event and shadow processing, apply to exclusive subsets of the input sonar data and may therefore occur in parallel; Modelling activity is based on the results of these analytic processes.

The set of subtasks comprising a complete interpreter is illustrated in Figure 2.6. The relationships between the tasks are also shown, both in terms of data and control interaction. In the following sections I shall introduce each module briefly and highlight the major research issues associated with its function. Two of the components, Viewpoint Registration and Specular Event Analysis, constitute the focus of the rest of the thesis.

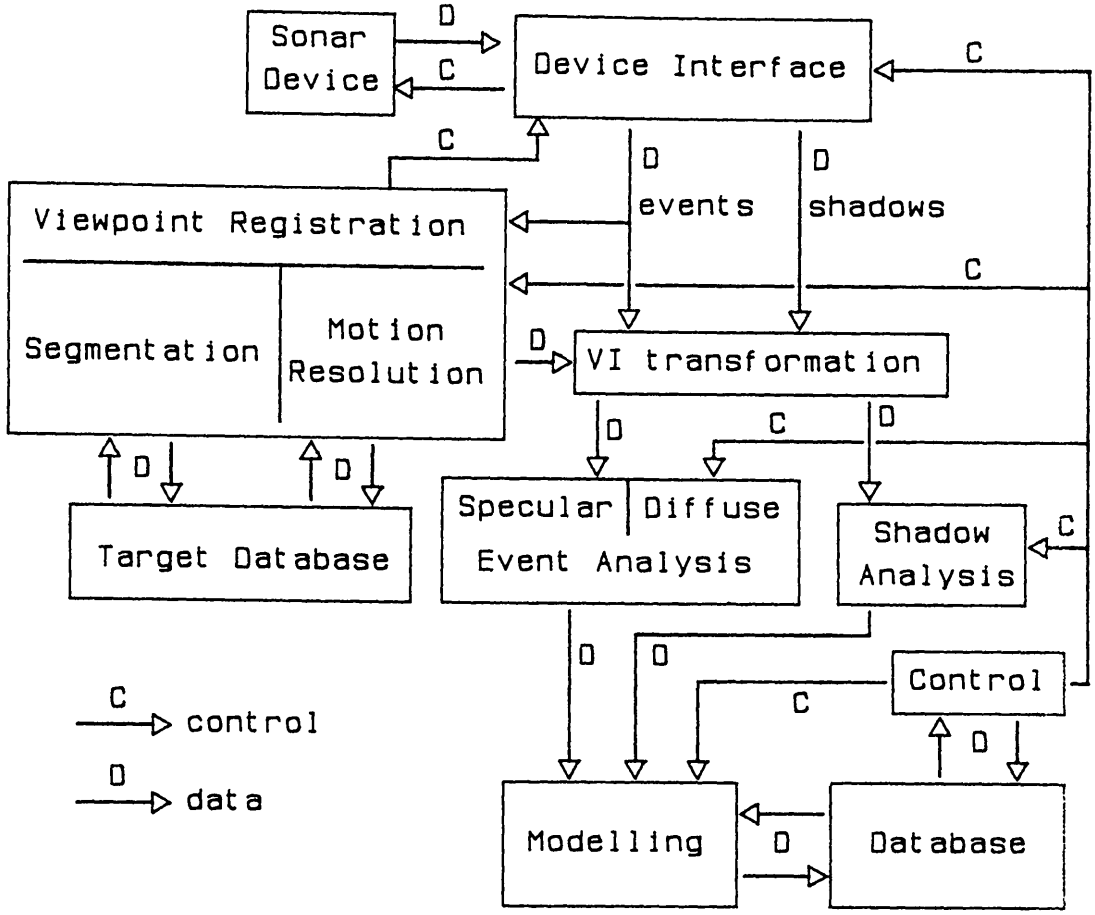


Figure 2.6. Organisation of a Complete Interpreter System.

2.2.1 The Sonar Device Interface.

This component of the interpreter has two functions: it performs necessary signal processing on the sonar video obtained from the driving sonar equipment; and it enables the middle and high level processes of the interpreter to be portable over a range of devices. It maps the physical sonar device into an abstract, virtual sonar.

There is a considerable variety of sonar systems suitable for driving sonar interpretation. The sonar may be coherent, able to estimate the Doppler velocity of targets, or incoherent, able to measure position only. It may be a mechanical sector scanning system, programmable on a per-scan basis; an electronically scanning phased-array system capable of irregular, fully programmable, scanning patterns; or a system of intermediate capability. It may be an anti-aliased system which can repeat measurements with a short

latency or an aliasing system with latency determined by its range limit.

The virtual sonar device assumed for the work presented later in this dissertation is a scanning sonar that provides a vector of information for each echo detection (each event). These event-vectors comprise the relative position of the detected echo with respect to the sonar transducer (represented in spherical polar coordinates) and certain other information may be included if it is available from the sonar. Examples of such optional data include the Doppler velocity of the echo source, echo strength, or spectral characteristics of the echo source. Event-vectors of this type are computable from a wide range of sonar systems, both commercially available sector scanning systems and state-of-the-art electronic beamforming systems, using standard signal processing techniques.

In addition to event-vectors the Sonar Device Interface produces descriptions of the shadows detected. These descriptions define the extent and location of the shadow relative to the observer, but the precise content of these shadow-vectors is an open research issue.

Finally, the Sonar Device Interface provides mechanisms for controlling the sensitivity and directionality of the physical sonar device.

2.2.2 Viewpoint Registration.

The event-vectors and shadow-vectors provided by the Sonar Device Interface describe echoes and shadows in the sonar reference frame which obtains at the time of their detection. The frame origin is the sensor position and the frame axes are aligned with the sensor range, azimuth and elevation axes (see Figure 2.7) which move as the vehicle orientation changes. The sensory reference frame changes continually because of vehicle motion. Viewpoint Registration is responsible for compensating for the continually varying viewpoint; it is fundamental to the viewpoint independent processing in the interpreter.

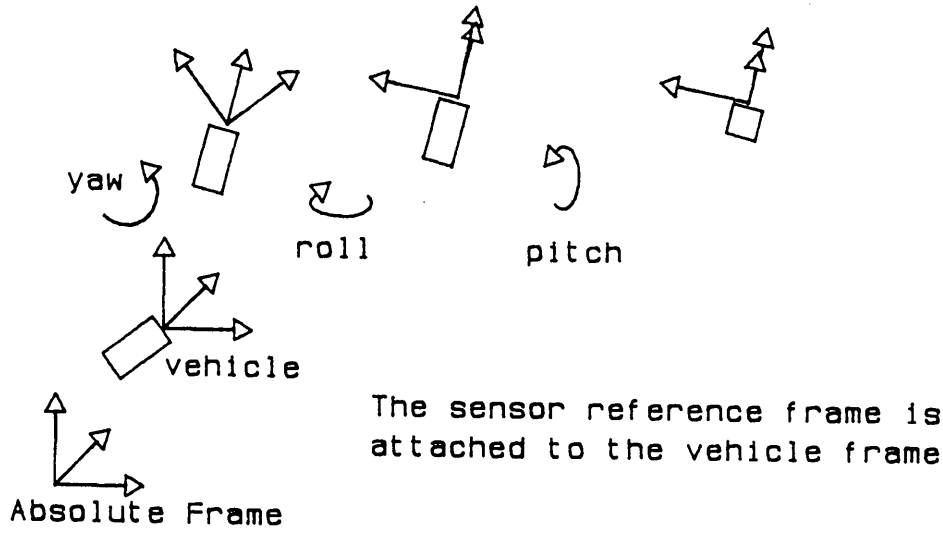


Figure 2.7. The Sensor Reference Frame.

There are two subtasks implicit in Viewpoint Registration: segmentation, and motion resolution. They interact using a shared data structure -- the Target Database.

The segmentation task attempts to collect events into groups that are generated by the same echo source. This first step of perceptual organisation consists in identifying consistent patterns of movement among the events, distributed in space and time, and postulating echo sources in the environment to account for those patterns. Each postulated echo source is noted in the Target Database and is associated with a sequence of events which are presumed to originate from that source. The echo sources recorded in the Target Database form the raw material for the Event Analysis described in the next section.

Each sequence of events generated during segmentation contains implicitly the relative position and velocity of its postulated echo source with respect to the sonar sensor. The motion resolution task resolves these relative motions into their components caused by sensor motion and by the proper motion of the echo source. The resolved sensor motion components are used to deduce the movement of the vehicle (navigational cues suplied by the vehicle control system are also used, if available) and thus the transformation between the sensor reference frame and a fixed stationary Cartesian reference frame is established. The echo source proper motion is recorded in the Target Database.

The result of the Viewpoint Registration process is twofold: a stable viewpoint independent reference frame is constructed by the motion resolution subtask as a side effect of computing the proper motion of the sensor, and a set of postulated echo sources that explain those events seen so far are created by the segmentation subtask. The description of each postulated echo source includes a history specifying which events in the input event sequence have been associated with the source, and an estimate of the position and proper motion of the source as computed by motion resolution.

The major research issues in the Viewpoint Registration problem are as follows: the solving of the motion resolution problem for continual observer motion in the presence of movement of environmental objects; the specification of the segmentation system; the strong real-time constraint on sonar interpretation, which applies more to this module (through which all the viewpoint dependent input data for interpretation must pass) than to any other component of the interpreter.

2.2.3 Event Analysis.

Each postulated echo source recorded in the Target Database by the Viewpoint Registration module is a generator of events. The reflections implied by those events may be specular or diffuse. In either case, the data associated with the echo source hypothesis constitutes a description of a part of an environmental object -- that part responsible for the events explained by the echo source hypothesis.

The Event Analysis module of the interpreter has the task of determining whether an echo source generates diffuse or specular reflections, and if the latter, determining what type of specular reflection is occurring. The raw material for these investigations is the description of the echo source and its motion history recorded in the Target Database.

Specular echoes arise by several mechanisms: from reflection in a concave corner; from reflection in a concave edge; from appropriate alignment of the observer and an object surface which may be planar,

cylindrically curved, or spherically curved. Each type of specular echo behaves in a characteristic way as the observer moves relative to the object. Specular Event Analysis has the task of determining the type of echo sources using movement histories of the source and observer. Once the echo source type has been identified, the geometric information in the echo source description may be interpreted in the light of the topological constraints entailed by the specular reflection mechanism in force for the source. The topology (loosely, the surface configuration and visibility properties) of the source are directly related to the local structure of the underlying object.

Diffuse reflection is not constrained in the same way as specular reflection and so does not provide topological cues. However, each diffuse reflection source corresponds to a point on the surface of an object and so constrains the geometry of that surface. Thus diffuse reflection source analysis provides additional geometric information for object modelling which is combined with the joint topological and geometric data of the specular sources.

The major research issues implicit in Event Analysis are first, the formulation of criteria for discriminating between diffuse echo sources and the various types of specular echo sources; second, the manipulation of uncertain evidence about the types of the echo sources, since only a small sample of all the possible views of the source will be available at any time; and third, the design of methods for incremental refinement of the source topology hypotheses as new events are associated with the source by the Viewpoint Registration module.

2.2.4 Shadow Analysis.

Shadow Analysis has the important job of processing the cues to object structure and size that acoustic shadows in typical sonar images present to the interpreter system. Shadows arise by occlusion, as in the familiar optical situation, but are perceived in terms of the ambient noise at the sonar sensor, as described above

(Figure 2.5).

The Shadow Analysis module is responsible for taking the descriptions of shadow regions constructed in the Sonar Device Interface and inferring the object surfaces that might cause them. The problem is complicated by the specularity of sonar, which means that the occluding surface may be acoustically invisible.

The major research issues inherent in Shadow Analysis appear to be the isolation of structure cues from the shadow regions and their relationships to the echoes (if any) returned from the occluding surfaces causing them, and the formulation of rules of inference that deduce object surface structure from shadow cues. This is an important area which, to my knowledge, has not been studied (although optical shadows -- where it is usually possible to see both the shadow and the surface causing it -- have been studied). However, a more detailed discussion is outwith the scope of this dissertation.

2.2.5 Modelling and Database Maintenance.

The highest (most organised) perceptual level of the interpreter system is object model construction and maintenance activity. The models constructed by the interpreter are *a posteriori* descriptions or explanations of the sensory input to the interpretation process. The modelling activity builds upon the results of Event and Shadow Analysis and is responsible for: postulating object surfaces and topology to account for observations analysed so far; updating surface and topological information as new evidence becomes available; enforcing consistency between various pieces of evidence available in support of a given object hypothesis; verifying that hypothesised models do indeed describe and explain the observations made so far; and, possibly, proposing strategies for movement or attention focus that help model construction or verification by obtaining necessary, but as yet unavailable, evidence.

The major research issues at this level are the precise definition of the model construction vocabulary and the extent to which the modelling activity directs the lower level activities of the

interpreter as opposed to being driven by them.

2.3 Two Facets of the Sonar Interpretation Problem.

Two components of the sonar interpreter modularisation proposed above were selected for deeper scrutiny. These were the Viewpoint Registration and Specular Event Analysis modules.

Viewpoint Registration was chosen because it is the fundamental activity of the interpreter system. For sonar interpretation, in the marine mobile robotic context, it is an essential prerequisite for the production of viewpoint independent descriptions of the environment. It is therefore of great intrinsic significance.

Viewpoint Registration is also of significance to land-based robot vehicles. The accepted paradigm for the navigation of land-based vehicles is the active motion 'Move; Stop and Think' paradigm exemplified in the work of Moravec with the cart robot (Moravec, 1980). A solution to the Viewpoint Registration problem based on a passive motion paradigm allows robotic vehicles to incorporate indirect motion information obtained from sensory processing into their navigation function in an essentially continuous manner.

A third motivation for focussing on the Viewpoint Registration module is that the solution of the novel, underconstrained, motion resolution problem opens the way to a full treatment of apparent environmental motion as well as observer motion.

The methods envisaged for the solution of the Viewpoint Registration problem motivated the second area of focus, the Specular Event Analysis module. The histories of observer and echo source position and proper motion computed by the Viewpoint Registration module make possible the discrimination of specular source topology. The topological and geometric information present in specular echoes makes them a rich channel of information for sensory interpretation.

A second motivation for this focus was its novelty. Although specularity has been studied in an optical context, it is hard to

discover the specular points of objects using optical sensors because optical reflection is relatively diffuse. With sonar sensors, discovering the specular points is much easier. Thus Specular Event Analysis of the type described here is immediately applicable only to sonar sensing.

Chapter 3. Sensor and Object Motion.

The task of computing sensor motion is a major problem in mobile robotic systems. In sonar interpretation it is a critical issue. A sonar transducer carried by a small marine vehicle is constantly in motion with respect to its neighbourhood and it is impractical, if not impossible, to keep the transducer stationary for an extended period. Furthermore, many objects in the marine environment have a proper motion and a sonar interpreter must be able to analyse such objects.

This chapter explores the implications of sensor and object motion in the context of marine sonar interpretation and defines the problem of motion resolution. It then discusses representative literature from the prolific work on motion extraction, compensation and analysis, and relates these efforts to the motion resolution task.

3.1 The Motion Resolution Problem.

The apparent motion of objects observed by a sensor (either acoustic or optical) has three component causes: the proper motion of the objects; the proper motion of the sensor (the observer); and noise arising in the imaging and transduction processes. Object proper motion is observed in a positive sense -- that is, the apparent motion component caused by the proper motion of the object has the same direction as that proper motion. The transducer motion contributes in a negative sense -- its component is in the opposite direction to the observer motion and for this reason I shall term it the reflected observer motion (or just reflected motion). Noise causing the location of objects to shift between observations gives rise to spurious random motion components. These effects are illustrated by the following example.

Let the observer be at the origin and suppose that there are two point objects, A at (4,0) and B at (2,8), with proper motions (0,-0.1) and (0.2,0.1) respectively. When the observer is stationary, the apparent motions of A and B are just these proper motions. This

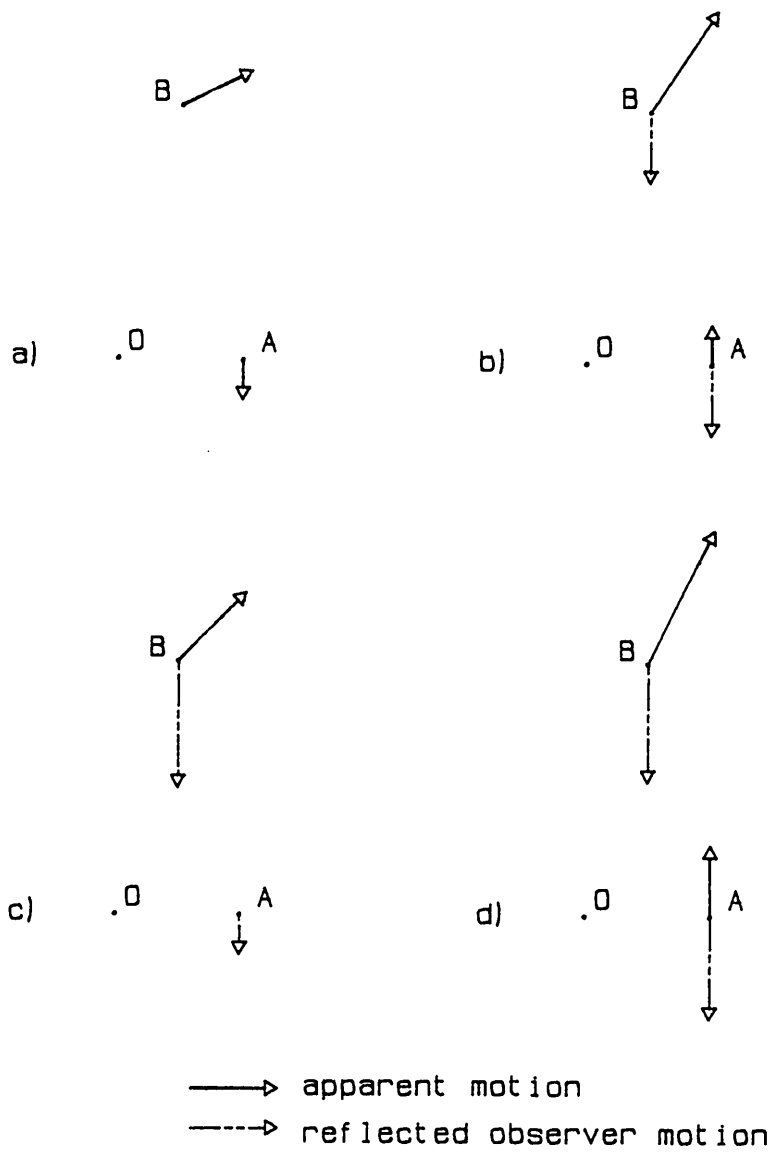


Figure 3.1. Components of Apparent Motion.

situation is shown in Figure 3.1a.

If the observer is moving the apparent motion contains a reflected motion component. In Figure 3.1b the observer is moving with velocity (0,-0.2) so the reflected motion component is (0,-0.2). This is subtracted from the object proper motions to give their apparent motions, which are thus (0,0.1) for A and (0.2,0.3) for B. Notice that the apparent motion of A is in the opposite direction to its proper motion in this example, because of the size of the reflected motion component.

An observer rotation also induces a reflected motion component. For an observer angular velocity of -0.025 radians per second, as in Figure 3.1c, the reflected motion is (0,-0.1) at A and (0.2,-0.05) at B, so the apparent motions in this case are (0,0) and (0,0.15) respectively. In this case A appears stationary because the observer motion cancels its proper motion.

Figure 3.1d illustrates the situation for an observer with both rotary and translational motion. The apparent motion is the vector difference of the proper motion and the reflected observer motion. If the object positions are perturbed by measurement noise the perturbations induce additive noise velocity vectors also.

The apparent motion observed is thus a composite of these three effects. Motion resolution is the task of separating out or "resolving" the various components of the observed apparent motion of objects. Equivalently, it is the task of deducing, using the apparent relative motion of each object, the proper (absolute) motion of that object and the proper motion of the observer. There are three major motivations for this task.

First, if the motion of the observer is known then it is possible to compensate the sensory input data for that motion, transforming the viewpoint dependent relative observations into a viewpoint independent frame of reference. Subsequent processing of the sensory input may then take place in the viewpoint independent reference frame. Viewpoint independent object models may be constructed, saving both the necessity of handling multiple representations of an object and the possibly time-consuming processing required to transform between those representations as the observer moves and the viewpoint changes. In the case of a continuously moving observer, such as obtains in the sonar interpreter context, the savings accrued by using a viewpoint independent frame are considerable.

Second, observer motion estimates derived by motion resolution from sensory data provide indirect observer motion information. This may be used to supplement or replace direct observer motion data (obtained from vehicle motion sensors) for control and navigation of the robot vehicle.

Third, the specularity inherent in the sonar imaging process makes relative motion between object and observer an important source of information. Different object configurations give rise to echoes that move with the observer in characteristic ways -- for instance, a

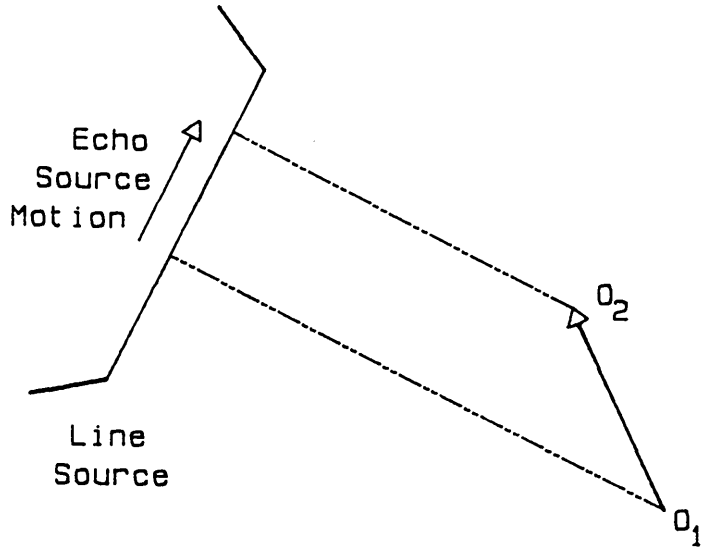


Figure 3.2. Reflected Motion from a Linear Reflector.

linear edge reflector (Figure 3.2) generates an echo that moves with that component of the observer velocity parallel to the edge. If both observer and echo motions are known with respect to a stationary absolute frame of reference (such as the viewpoint independent frame discussed above) then the object configurations (the source topologies) generating moving and stationary echoes may be deduced.

We have been discussing motion resolution in terms of object proper motion. However, the specularity of sonar imaging ensures that complete object images are rarely seen -- rather each object is observed as a number of echo sources. To talk of objects thus implies a knowledge of the relationship between echo sources and the objects responsible for them; for sonar interpretation this is begging the question -- the construction of such an association constitutes part of an "interpretation" of the sensory data. While it is possible to use the information contained in a priori or partially constructed object models to control motion resolution by constraining the relative and absolute motions of echo sources, and hence their proper motion components, it is inappropriate to use this global high-level data at such a fundamental level of the interpreter. I therefore choose to formulate motion resolution in an object-

model-free way, in terms of the motion of echo sources rather than object motion.

> The motion resolution problem is the task of resolving the observed composite motion for each echo source into its proper motion and reflected observer motion components and of deducing from these components the absolute proper motions of observer and echo sources.

3.1.1 Solution Design Constraints.

The motion resolution problem is of very broad application; as I shall argue later (in section 3.2.1) any robotic system that makes use of sensory data must address the problem for its particular sensory environment. For sonar interpretation this context is the marine acoustic imaging process, and the interaction of the properties of marine sonar and the demands of the motion resolution problem give rise to various constraints on the motion resolution solution. These constraints are enumerated below.

The first three constraints were presented in Chapter two as applying to the sonar interpreter system as a whole.

C1 The input data for motion resolution processing is viewpoint dependent, comprising relative position vectors of echo sources with respect to the observer.

C2 Echo sources are only intermittently (randomly) visible to the sonar.

C3 Any solution to motion resolution for sonar interpretation must be able to compensate for a continually changing observer viewpoint.

Viewpoint Registration implies a segmentation problem -- the event vectors provided by the sonar device must be matched to the echo sources that cause them -- which may be solved more easily if the motion resolution solution can provide predictions of the absolute and relative motions of echo sources, and can indicate the accuracy of such predictions.

C4 Any motion resolution solution must be able to estimate the position and velocity of observer or echo source at any time (including future times) and provide a confidence interval for that estimate.

To these design constraints we may add one implementation constraint:

C5 The motion resolution solution must be implementable as a part of a real-time sonar interpretation system.

3.2 History of Motion Resolution.

The motion resolution problem has been stated in terms of observer and echo source motion components to be resolved from the composite echo source motions apparent to the sonar sensor. Equivalent more general formulations of the problem for an unspecified sense are:

> compute the proper motion of a robotic vehicle from its relative motion with respect to the stationary objects in the world observed by its sensors,

and

> construct and maintain a viewpoint independent frame of reference with respect to which the motion of the robot and of objects in its world is deduced using relative motion measurements obtained from a sensory system.

In this section, the relevance to the motion resolution problem of work in the areas of robot navigation, terrain-aided navigation, optical flow, and target tracking will be explored. Each of the methods presented is inadequate for the sonar interpretation application; the deficiencies of each approach will be highlighted and discussed.

Many of the approaches we shall consider may be described as optically orientated in that they presuppose a contiguous input array of pixel data -- intensity, range, or optical flow velocity vector values. This optical orientation limits their relevance to sonar interpretation for two related reasons.

First, because of the inherent specularity in the sonar imaging process, the distribution of information in sonar data differs from that in optically orientated images. The former consists of a sparse collection of bright reflections (perhaps several hundred events), each having a high information content, whereas the latter consists of a large array (typically several thousand elements) of highly redundant pixels.

This radically different distribution of information requires different processing techniques. Whilst it is possible to collect optically orientated images from sonar (using acoustic imaging techniques, for example (Sutton, 1979)) for optically orientated processing, to do so neglects the high information content of the relatively few specular reflections in the sonar image. These reflections provide not only the range and bearings to a point on an echo source but also give some information about the surface normal at that point. To collect an image and then process it to obtain this information is a waste of effort.

Secondly, the optically orientated approaches are batch style processes -- they operate on large quantities of input data simultaneously, often taking considerable time. For sonar interpretation, with its strong real-time performance constraint, the use of recursive[1] techniques on the relatively few events in an acoustic range

1. The term "recursive" is used here in the normal sense of statistical filtering, i.e. a process that computes its output on the basis of its current input and previous outputs.

image represents a better use of processing resources.

3.2.1 Robot Navigation.

Motion resolution affects all robotic systems to some extent. In its simplest form it must be solved once and for all by a robot which intends using a sensory system for feedback, in order to establish the relationship between the sensor reference frame and the (viewpoint independent) world or robot reference frame. This calibration process may be seen in, for example, Popplestone and Ambler's striper system (Popplestone and Ambler, 1977) where a camera calibration matrix computed at the start of a run defines the relationship between camera and world coordinates. In effect, the world coordinate frame has been established by defining its relationship to the sensor proper reference frame.

Camera model identification plays an important role in the ACRONYM system also (Brooks, 1981). In this case, however, the model identification is distributed over time as the system processes object model constraints; the model may be underdetermined for all or part of the processing done by ACRONYM. The final model identification fixing the relationship between camera and world reference frames need not occur, since ACRONYM can reason with partially constrained relationships, but in that case the positions of objects in its field of view may not be fully determinate either.

This once-for-all approach works in these cases because the sensors are stationary with respect to the world. Clearly it will work in any situation where the relationship between the sensor reference frames at different times is easily known.

In the case of mobile robotic devices, the problem is complicated somewhat. If the robot mechanism is accurate and repeatable, a once-for-all sensor calibration may suffice; for vehicles that are not particularly accurate or must cope with unexpected perturbations, the reference frame established by a once-for-all calibration soon loses its alignment with the world. Under these circumstances a solution is to recalibrate periodically.

There are two ways to recalibrate the robot once it is decided that the current reference frame transformation is no longer accurate. Either the vehicle moves to a known and recognisable place (or uses known, recognisable features in the environment) and the reference frame transformation is reset using stored calibration values; or the navigation system attempts to compute how the vehicle has moved since its last calibration, using sensory data alone, and when this motion is known the new reference frame transformation is computed from the old.

I shall refer to these two approaches as model-dependent and model-independent respectively. In the former, the robot is supplied with a map or model of its environment, specifying readily recognisable features for use in recalibration -- the model provides the necessary data to determine the current position of the vehicle from an initial estimate of its position and its current view of the local world. Typically, this style of navigation is used in controlled environments. Examples include the factory vehicles being developed by Larcombe (1981), which contain a compiled model of their section of factory and use acoustic sensors, and the building inspector robot of Yachida, Ichinose and Tsuji (1983), which possesses a model of its building and uses optical sensors.

The building inspector vehicle of Yachida et al. uses an environmental model for the dual purpose of navigating and monitoring changes in the building. Initial calibration of the camera is achieved using a known object in a standard position relative to the vehicle. Then the vehicle recalibrates its position every few metres along its route using the current image of its environment. To do this, first the camera pan angle is estimated accurately using the position in the image of the floor vanishing point. Once the pan angle is known, vertical lines in the image are matched with verticals predicted by the building model -- an approximate vehicle position is available from position sensors in the robot's drive system and from the last calibration position. The correspondence between image and model verticals is used to compute an accurate position estimate which is used as the new calibration position. Vertical lines are chosen as the features to use in the recalibration because

they are plentiful and stable features in images of the environment.

This model-dependent approach is satisfactory in controlled and relatively predictable environments such as building interiors. In situations where suitable models are not known or where the world is unpredictable, such as outdoor terrain applications, a model-independent technique that does not require extensive a priori knowledge is appropriate.

This methodology is exemplified in work by Moravec (1979, 1980) on the Stanford cart robot. Moravec addressed the problem of navigating accurately over outdoor terrain. The vehicle was equipped with rough displacement sensing using encoders attached to its wheels and with a television camera mounted on a linear track. Nine images, collected by the camera at different places along its track, were filtered using a special "interest operator" to detect high contrast regions in the image; a disparity algorithm then combined the nine sets of image points, computing the three-dimensional position of each "interesting" world feature with respect to the camera. The set of chosen features, fixed in the world frame of reference, effectively define the transformation between the camera (and robot) reference frame and the world frame.

Once this calibration is complete the robot executes a roughly known planned movement -- the movement is not known exactly because of wheel slip, uneven terrain, drive errors and so on. A new set of visual input is then collected. Reference features common to this set and the visual data obtained before the movement are used to compute a least square error transformation between the camera positions before and after the movement (using the method developed by Gennery (1977)). The rough displacement from the wheel sensors provides a starting point for this estimation. The recalibration takes some time and until it is complete the vehicle cannot relate its world reference frames before and after the movement.

From the standpoint of generality, this approach is superior to the model-dependent one. It does not require a priori knowledge of the world and relies for navigation on features which exist in all (except pathologically specialised) environments. It is the most

promising of the two methods for the sonar interpreter application, but unfortunately it falls short of the mark for two reasons.

The essential deficiency with Moravec's model-independent approach is that it is discrete. The strategy for navigation can be described by the epithet 'Move; Stop and Think' (MSaT). The vehicle moves, then remains stationary for a time while it recalibrates in preparation for its next move. In the marine world, where the vehicle cannot be held still, this paradigm founders precisely because it needs to stop and think. In the case of the cart the difficulty is in processing the nine input images and computing the least square error estimate of the six-parameter transformation between camera reference frames. With special purpose hardware, or different sensory arrangements, this extended and demanding processing could be attenuated to the point where the vehicle spent relatively little time stationary.

Such hardware optimisation at first sight appears to overcome the problems of relating the 'Move; Stop and Think' paradigm to the continuous motion in the marine context. There remains, however, a fundamental philosophical objection to the approach -- a vehicle adopting the MSaT strategy of navigation tacitly assumes that it is the agent of its own motion. This was appropriate for Moravec's work with the cart and for Yachida et al.'s building inspector (which uses a model-dependent MSaT strategy) but is inappropriate for a marine vehicle which, although powered and under control, suffers continual random and systematic perturbation of its motion by marine currents. Such a vehicle is clearly not the agent of all of its motion (though, of course, it is in control of the overall average motion); motion is inescapably forced upon it by its environment. The 'Move; Stop and Think' paradigm does not address this problem of continual unplanned movement.

Collision Avoidance using Acoustic Sensors.

Another possibility for using acoustic sensory data in robot navigation is exemplified by the HILAIRE vehicle, which carries 14 short range acoustic sensors that provide warning of unexpected obstacles

and are used by the vehicle control system in wall following and other simple closed loop navigation tasks (Bauzil, Briot and Riges, 1981). However, the main sense mode for the HILAIRE robot is visual.

Work by Lane (1984) with the ANGUS vehicle has a similar emphasis on automatic guidance of the vehicle in simple navigation tasks using sonar data. A "blocks world" sonar simulation system has been constructed, able to generate crude but realistic sonar images from models of the structures present in the marine environment. Using this simulator, which can be coupled to the vehicle guidance simulation, it is intended to develop a closed loop collision avoidance reflex which will then be tested using the ANGUS vehicle and real sonar data. The raw sonar data is processed using image enhancement techniques and targets of interest to the vehicle guidance system will be selected and analysed using a small expert system.

Since these projects are concerned with using acoustic data for collision avoidance rather than navigation, their relevance to the motion resolution problem is limited. However, Lane's work is relevant to other aspects of the sonar interpreter design, notably the Sonar Device Interface.

3.2.2 Terrain-Aided Navigation.

The navigation techniques used with the Stanford cart and the building inspector vehicles could be described as terrain-aided navigation systems. Sensory data is used to improve the accuracy and repeatability of the primitive position measuring systems built into the robot drive. The term 'terrain-aided navigation' is more conventionally used to describe methods for using sensory data obtained either optically or by radar to correct drift and other errors in the inertial navigation system (INS) of an aircraft or missile. In this section we shall consider three such systems, two radar driven and one optically based.

The terrain contour matching (TERCOM) navigation system, devised for Cruise missiles, uses a terrain profile derived from radar altimetry to improve estimates of position, velocity and other states

of the missile INS (e.g. Andreas, Hostetler and Beckman, 1978; Reed and Hogan, 1979). The ground clearance profile obtained from the sensor is correlated with a stored terrain map, and the position of maximal correlation is used as an input to a Kalman tracking filter estimating the INS state. The technique is a batch process (this is clear in (Reed and Hogan, 1979), for example) because the input data has to be correlated with the map, and it is model-dependent -- the model is the a priori map of terrain topography.

An alternative realisation of the TERCOM idea, proposed by Hostetler and Andreas (1983), is model-dependent and recursive. Raw ground-clearance measurements from the radar altimeter are filtered directly using a Kalman filter that incorporates a linearisation of the local terrain. Two linearisations were evaluated in (Hostetler and Andreas, 1983): a simple first order derivative of the local ground surface; and a stochastic planar fit to a local terrain neighbourhood. The latter method computed a least square error plane approximation to a neighbourhood whose size depended on the current position uncertainty. The error in the fit contributed to the filter noise model. The former algorithm was found to fail if the initial position uncertainty was moderately large (about 75 metres standard deviation), with the filter state diverging because of the highly non-linear nature of the problem. The latter algorithm, being able to use the low frequency characteristics of the terrain when position uncertainty is large, and to focus on smaller, more detailed, terrain regions as the estimates improve, did not suffer from divergence until initial position uncertainty became very large. However, for larger initial position uncertainties, the convergence of the filter is appreciably delayed by the smoothing of the terrain implicit in the stochastic linearisation.

The technique suggested by Hostetler and Andreas to overcome the convergence delay is to use a number of filters initialised at different points in the initial uncertainty region. The filters are run in parallel and the one with smallest averaged weighted residuals[1] is chosen as the correct one. A similar multiple model estimation

1. that is, the running average of filter innovation weighted using the innovation covariance.

algorithm is evaluated by Mealy and Tang (1983) for helicopter navigation. In that work the a posteriori probability of correctness is used as the filter selection criterion. For both algorithms, once a filter is chosen as correct the multiple model estimation is collapsed to a single filter.

The third method of terrain-aided navigation considered here is Hannah's "Bootstrap Stereo" technique (Hannah, 1980). The method is optically driven, using a television camera sensor, and model-independent; it is a slightly modified version of Moravec's 'Move; Stop and Think' navigation scheme for the cart, and forms a part of an aircraft "Navigation Expert". The Bootstrap Stereo algorithm

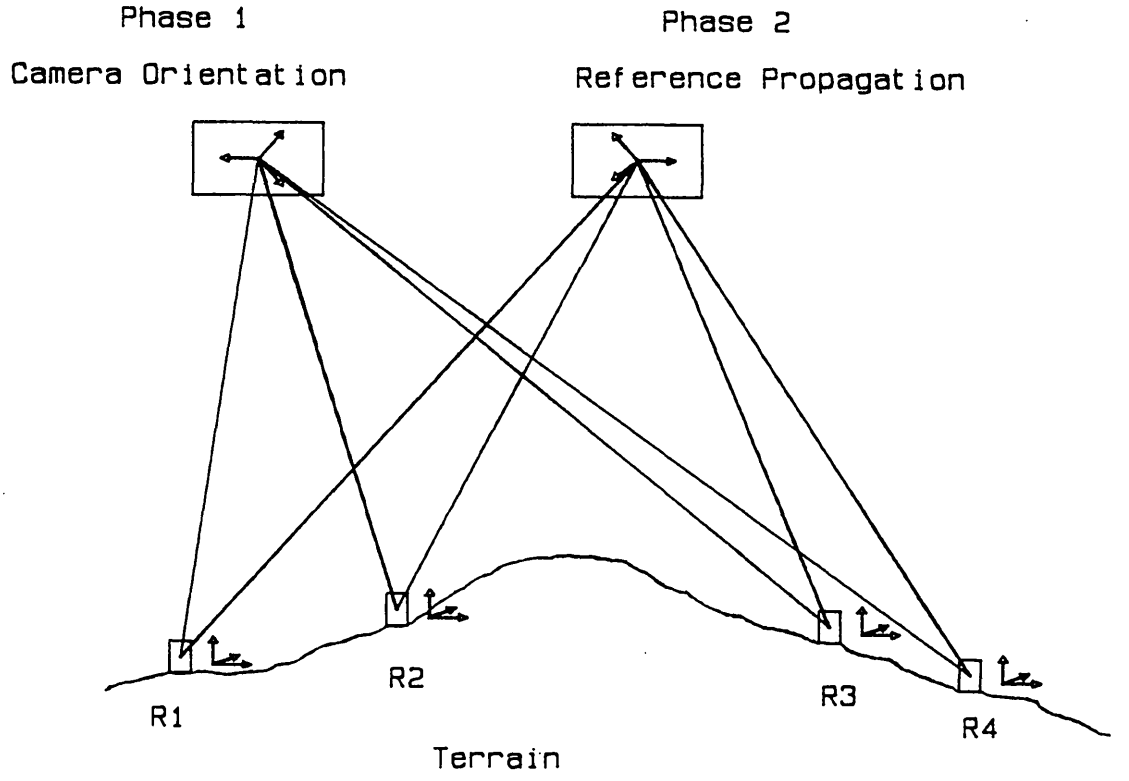


Figure 3.3. Operation of Bootstrap Stereo Navigation.

iterates two phases as depicted in Figure 3.3: given the positions of known points on the ground, the camera position and orientation are determined from the image positions of the points (phase one); and given the camera position at two times, the position of points on the ground is determined by triangulation (phase two). The algorithm is bootstrapped by specifying the position of a few points, and continues to function by computing the positions of new points in phase two

which then become given points for the following phase one. Thus the method requires a small input of a priori information, but for the most part functions without any model of the terrain.

None of these methods is particularly suitable for the motion resolution problem. The TERCOM techniques require an accurate and detailed a priori model of the terrain over which the vehicle will navigate. Bootstrap Stereo is model-independent, but it does not address the issues of moving reference points and automatic initial reference point selection.

3.2.3 Optical Flow.

Recent years have seen a resurgence of interest in the use and interpretation of optical flow fields. First described by Helmholtz (1925), such fields arise because of the relative motion of an observer with respect to his environment, which causes the positions and directions of image features to vary with time and so defines a vector field of velocities within the image of the environment.

Optical flow field interpretation and motion resolution are closely related problems. Both are based on the motion of an observer with respect to his environment, and optical flow potentially affords a solution to the continual motion inherent in the motion resolution problem we have formulated. As its name suggests, however, it is optically orientated. In this section we shall consider in detail the relationship between optical flow and reflected motion, the use to which each may be put, and show that the motion resolution problem is at once less demanding and further reaching than the interpretation of optical flow.

Consider the situation depicted in Figure 3.4. The observer, at O, is moving (with respect to some absolute frame of reference) with velocity $\underline{v}_O$ and angular velocity $\underline{\omega}_O$; an object point P is moving with velocity $\underline{v}_P$. The perceived motion of P at O is just the relative motion of P with respect to O, projected by whatever imaging system the observer is using. Now the relative velocity of P with respect to O, ${}_P\underline{v}_O$, is given by

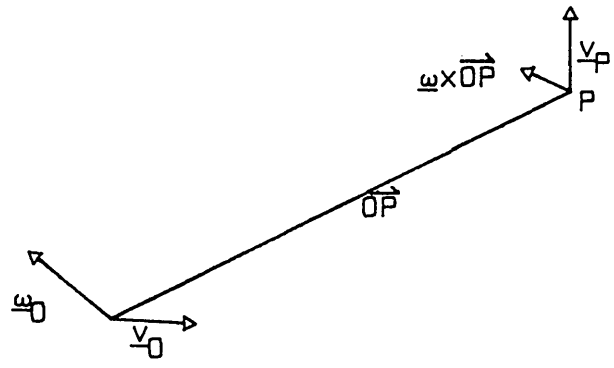


Figure 3.4. Relative Motion of an Object Point.

$$ {}_P\underline{v}_0 = \underline{v}_P - \underline{v}_0 - \underline{\omega}_0 \times \overrightarrow{OP} \qquad (3.1) $$

since the velocities of O and P are absolute velocities. In this situation the motion resolution problem is to resolve the composite velocity ${}_P\underline{v}_0$ into its component due to the object absolute motion $\underline{v}_P$ and the reflected motion component

$$ \underline{rm}(\overrightarrow{OP}) = \underline{v}_0 + \underline{\omega}_0 \times \overrightarrow{OP}. \qquad (3.2) $$

From its definition in equation (3.2) it is clear that the reflected motion $\underline{rm}(\underline{r})$ is a vector field comprising two parts -- a position-independent translational component $\underline{v}_0$ and a position-dependent rotational component $\underline{\omega}_0 \times \underline{r}$. It is a three-dimensional vector field and is defined at every point of the Euclidean 3-space E^3.

The optical flow field is derived from the apparent motion field by projection. Figure 3.5 illustrates the relationship, making explicit the imaging process used by the observer. Figure 3.5a depicts planar perspective projection, an example of which is the imaging scheme of a conventional camera; Figure 3.5b shows projection onto a hemispherical imaging surface, which is often used to model imaging in natural eyes. In either case the apparent motion is projected onto the imaging surface, giving rise to a two-dimensional vector field in that surface -- the optical flow field.

The optical flow field, like the apparent motion field, comprises a translational and a rotational component. If all the objects in the environment are stationary, then the translational field radiates

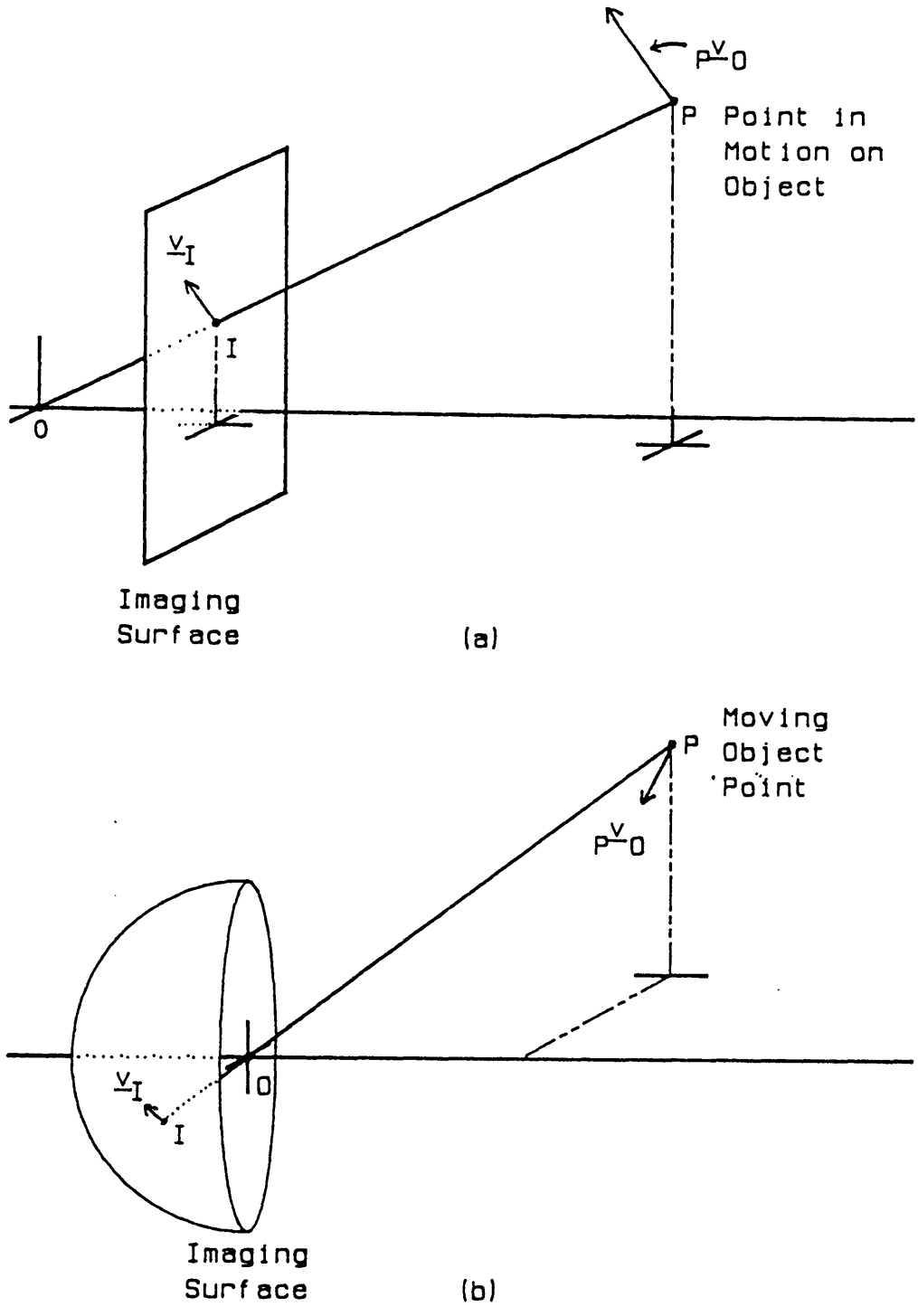


Figure 3.5. Optical Flow caused by Apparent Motion.

towards or away from a focus of expansion (or vanishing point) that is the intersection of the observer's linear velocity with the imaging surface. The rotational field is directed around a pole where the observer angular velocity vector intersects the imaging surface. The translational field at an image point depends on the depth and

direction of the corresponding object point and the observer's linear velocity, but not on the observer's angular velocity, while the rotational field depends only on the direction of the object point and the angular velocity of the observer.

If objects may move, these properties still hold for the optical flow field over the image of any given object, provided that the observer linear velocity is measured with respect to that object. A summary of the characteristics of the apparent motion and optical flow fields is given in Table 3.1.

Apparent Motion	Optical Flow
3D Vector Field	2D Vector Field
Position-independent Translational part	Depth-dependent Translational part
Position-dependent Rotational part	Depth-independent Rotational part

Table 3.1. Comparison of Reflected Motion and Optical Flow.

The optical flow field is a rich source of environmental information, containing topological information on object structure (Koenderink and van Doorn, 1977), and can be inverted to provide orientation of object surfaces (Clocksin, 1978), relative depth, and observer motion (Prazdny, 1980). Longuet-Higgins and Prazdny (1980) showed that at any point where there is a depth discontinuity in the image the relative depth and full observer parameters (the linear and angular velocity) can be obtained on the assumption that the world is rigid. If the optical flow and its two-dimensional spatial derivatives are known for a non-planar surface patch then the relative motion of the observer and surface patch can be found. Rieger and Lawton (1983) showed how to use the translational flow field to deduce motion parameters for curvilinear navigation.

The analysis and interpretation of optical flow fields, therefore, is concerned with the inversion of the observer's imaging projection to recover the apparent motion field, object surface characteristics and object topology. Motion resolution, on the other hand, concerns

the interpretation of the apparent motion field. The observer motion deduced from optical flow is the relative motion of observer and an object; motion resolution processes such relative motion data and yields absolute motion information. Thus the motion resolution problem is less demanding than optical flow analysis in that it presupposes the apparent motion field rather than its projection (and so does not have a singular mapping to "invert"), but is further reaching since it goes beyond optical flow, computing absolute motions and handling environments containing moving objects.

3.2.4 Target Tracking Systems for Radar and Sonar.

The last section reviewed the processing possible when an optical flow field is available to a vision system. In the context of sonar interpretation, such processing is inappropriate to motion resolution because it is possible to estimate the apparent motion field directly, sidestepping the necessity of computing and inverting its retinal projection. We therefore turn our attention in this section to methods that allow the apparent motion of points to be estimated from sequences of relative position vectors -- the techniques extensively employed in radar and sonar signal processing systems.

Just as tracking features in image sequences implies a correspondence problem, so also the tracking of anonymous radar or sonar target sightings implies a correspondence problem. The tracking system must deduce which known (or new, unknown) target caused each detection and hence which track is to be updated using the new relative position, or else dismiss the detection as noise (clutter). This problem, which I have termed the segmentation problem, has been extensively treated in the literature of tracking systems and is discussed in detail in Chapter seven. We shall assume here that this correspondence problem can and has been solved.

The importance in radar and sonar applications of tracking moving targets and eliminating stationary "clutter" and false detections has resulted in a proliferation of algorithms for direct object tracking; a great deal of effort and ingenuity has also been invested in

developing systems that run in real time. Since the work of Kalman in 1960 the methods for computing smoothed motion estimates for targets and predicting future target positions have all been recursive filters of varying complexity, generally the α-β and (ordinary and extended) Kalman filters.

In radar applications, where the need is for high speed processing of large numbers of detections and where models of the target motion can be simple, the α-β and the Kalman filter are generally used (e.g. Morley and Wilsdon, 1977; Holmes, 1977; Buchner, 1977). Both are examples of linear statistical estimation filters, operating on a sequence of noisy measurements to compute a corresponding sequence of estimated quantities. They employ the cyclic computation scheme illustrated in Figure 3.6.

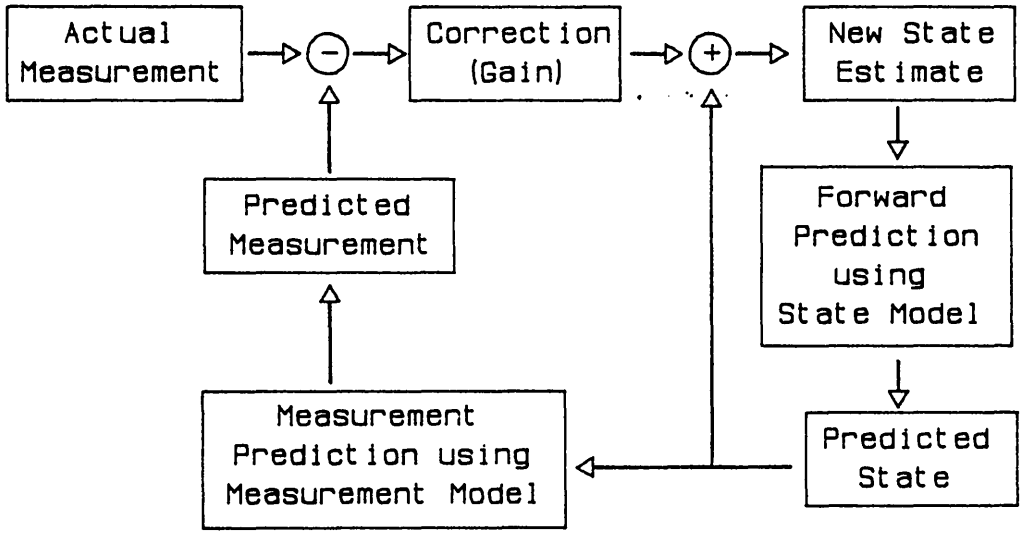


Figure 3.6. Operation of Linear Estimation Filters.

When a new measurement, x_n, arrives, the filter computes a predicted value for that measurement using a prediction, $\underline{\tilde{x}}_n$, of its state vector for the time when the new data is due. The error difference between the prediction and measurement is multiplied by the filter gain to give a correction term which is added to the state prediction, giving a new state estimate $\underline{\hat{x}}_n$.

The α-β filter estimates a quantity (x, say) and its rate of change (v). Its state vector comprises estimates of x and v, and the filter gain is a corresponding two component vector comprising the position error gain α and the velocity error gain $\frac{\beta}{\tau}$, where τ is the constant intersample time of the measurement sequence. The filter has no explicit measurement model -- the predicted measurement is just the current prediction of x -- and the state transition model is simple: the prediction of x is $\hat{x} + \hat{v}\tau$, where $\hat{x}$ and $\hat{v}$ are the current estimates of x and v; the predicted value of v is just $\hat{v}$.

This filter requires a constant intersample time sequence of input measurements with constant accuracy. In its simplest form it has constant gain, but it may easily be extended to provide a linear regression filter by allowing the gain parameters α and β to vary with time (Marks, 1961). Although essentially a scalar filter, it may be ganged to filter vector quantities with uncorrelated components. It is economical both in computation and in storage.

The Kalman filter is a powerful general purpose linear filter. It incorporates explicit mathematical models of the measurement process and of the filter state variation with time, so it is able to accommodate an input sequence with variable intersample time comprising data of varying accuracy and differing relationships to the filter state vector. It is a true vector filter, allowing cross-correlation between the errors in measurement state estimate vector components. The state estimate outputs are accompanied by estimates of the output error covariances, and the filter gain is chosen for each cycle to minimise the output error covariance.

The Kalman filter is optimal (it produces minimal covariance outputs) for unbiassed unimodal error distributions. It can be costly to compute and requires more storage than the α-β filter. For applications where the state model is non-linear the filter may be readily enhanced to use a slowly varying linearised version of the non-linear model -- this enhancement is called the "Extended Kalman filter".

Examples of Kalman and α-β filters in use abound in the tracking literature. Morley and Wilsdon (1977) describe a multiradar tracking system in which signals from individual radars are processed using

least squares α-β filters and the results combined to derive target track estimates using a Kalman filter. Holmes (1977) uses α-β filters as the basis of his automatic track initiation scheme but modifies them to allow successful recovery from target fading by altering the gain parameter when a fade is detected. He argues from experience that the α-β filter modified in this way is a good replacement for uncoupled Kalman filters (where there is no correlation between the measurement or state dimensions) with a substantial reduction in storage and computational load. Buchner (1977) uses Kalman filtering to track targets using a variety of types of measurement (range sum, range difference, and bearing) derived from a number of radar sites. For passive sonar applications, where the data rate is lower, the estimation task is non-linear and the number of state parameters to be estimated is large, extended Kalman filters are frequently used (for example, Fortmann, Bar-Shalom and Scheffé, 1980, 1983).

There are two main problems in the use of Kalman filters -- bias in the input measurements, and filter divergence caused by numerical problems.

The filters are sub-optimal in the presence of bias, but correction can be made for the bias by augmenting the filter state so that the bias values are also estimated (Friedland, 1969), or by filtering the innovations to extract the bias value and compensating the input measurements (Russell and Bugge (1981) take this approach; they also use the innovation sequence variance to set the filter gain).

If calculated naively the Kalman filter covariance update formulae are $O(n^3)$ in complexity and prone to numerical instability, which manifests itself in erroneous covariance and gain matrices. These may differ by several orders of magnitude from the true errors in the filter state, and cause divergence of the system and subsequent tracking loss. The problems may be tackled together using matrix factorisation techniques such as Cholesky or U-D decomposition, or by calculating covariance and gain schedules off-line or using closed form solutions to the Ricatti equation. For example, Thornton and Bierman (1978) present an efficient U-D realisation of the Kalman

filter and demonstrate the substantial improvement in the accuracy of the computed covariances achieved by this technique. Gupta and Ahn (1983), Orfanidis (1982), and Fitzgerald (1981), construct closed solutions to the Ricatti equation for certain simple filters of the type used in radar applications, for constant intersample time and data accuracy, which can then be used to calculate the covariance and gain for the steady state filter (after the startup transients have died out). The precalculation or closed form techniques do, however, require a large amount of a priori knowledge of the measurements.

The α-β and Kalman filters, then, can be used to estimate the apparent motion of targets given only their observer-relative positions at various times. In the sonar context, where targets are intermittently visible (so that the time between measurements is widely variable) the Kalman filter is indicated despite its extra storage requirement and complexity; the error covariance estimates generated by that filter simplify the segmentation processing and enable motion resolution processing to propagate error estimates of the absolute motion estimates and to attempt to make optimal use of the information present in the apparent motion. The Kalman filter simplifies greatly in certain special cases, and, by careful choice of coordinate system and the use of matrix factorisations, large savings can be made. This point will be discussed further in Chapters five and six.

3.2.5 Summary.

A solution to the motion resolution problem must deduce absolute motions from a sparse, noisy, apparent motion field in an efficient, model-independent, way. It must adequately treat both active and passive observer motion, handle object motion, and preferably be recursive in processing style. The work we have been considering fails to meet these criteria: model-independent navigation suffers from an inadequate motion paradigm; terrain-aided navigation is model-dependent; optical flow techniques and target tracking methods are concerned with the computation of relative motion (though by assuming a static environment some are able to calculate the observer

proper motion) which is the starting point for motion resolution.

Chapter 4. Motion Resolution for Linear Observer Movement.

Motion resolution is the problem of separating or resolving the composite apparent motion of echo sources relative to an observer into the component due to the observer proper motion and that caused by the motion of the echo sources themselves. It is a crucial issue for a mobile robotic system navigating in an environment containing moving objects and unable fully to control its own motion. It is a part of the Viewpoint Registration processing essential to sonar interpretation.

In this chapter I present a solution to the motion resolution problem suitable for situations where the observer moves only linearly. The method is presented and evaluated in terms of two-dimensional motion but, because it is algebraically linear, it extends immediately to the three-dimensional linear motion case.

4.1 The Linear Motion Resolution Problem.

The general motion resolution problem permits full three-dimensional observer movement. Such motion has six degrees of freedom -- three translational ones corresponding to a changing observer position, and three rotational ones corresponding to a changing observer orientation. The work presented in this chapter restricts this general problem in two ways.

First, it assumes that the only observer motion to be determined by the motion resolution system is translational. This case obtains if the observer's sensor axes are fixed with respect to the world or, more generally, if the orientation of the sonar axes can be determined directly *a priori*. In a situation where only translational motion of the observer is allowed the motion resolution problem will be termed *linear*.

The second restriction made here is to consider only two-dimensional observer motion. This does not affect the applicability of the solution to the three-dimensional case: the solution is

defined in a way that makes no explicit reference to the dimensionality of the problem and is thus immediately extensible to three-dimensional motions.

The two-dimensional linear motion resolution problem represents a considerable simplification of the general problem, yet it is a reasonable simplification. The linear problem retains the significant features of the general problem (the necessity of handling both active and passive observer motion and of recovering object proper motion from apparent motion) while avoiding the complexity of three-dimensional rotary motion. The extension to handle full three-dimensional motion is the subject of Chapter six.

Linear motion resolution is also a useful tool in its own right. It may be used by any robotic vehicle which makes only planar motions (for example, those robots that move on the floors of buildings) and which has means of determining its orientation independently of the motion resolution processing. (This latter condition is relaxed for two-dimensional motion in Chapter five.)

4.1.1 The Indeterminacy of Motion Resolution.

In this section I shall analyse the principal difficulty of motion resolution -- its implicit circularity. The starting point for motion resolution is a knowledge of target relative position and apparent motions, and its goal is the extraction of the observer and target absolute motions (motions with respect to a stationary viewpoint independent frame). If the observer and target absolute positions are $\underline{p}_O$ and $\underline{p}_P$ respectively, and the target proper motion is $\underline{v}_P$, then the relative position, $\overrightarrow{OP}$, and apparent motion, ${}_P\underline{v}_O$, are given by equation

$$\overrightarrow{OP} = \underline{p}_P - \underline{p}_O, \tag{4.1a}$$

$${}_P\underline{v}_O = \underline{v}_P - \underline{rm}(\overrightarrow{OP}), \tag{4.1b}$$

where $\underline{rm}$ is the reflected motion component defined by equation (3.2), reproduced here:

$$\underline{rm}(\underline{r}) = \underline{v}_0 + \underline{\omega}_0 \times \underline{r}. \qquad (3.2)$$

Without loss of generality $\underline{\omega}_0$ may be set to zero for the linear motion resolution problem (if the observer does in fact rotate the observations may be corrected for this using a priori knowledge of the rotation). These relationships, which I shall call the equations of observation, are illustrated in Figure 4.1.

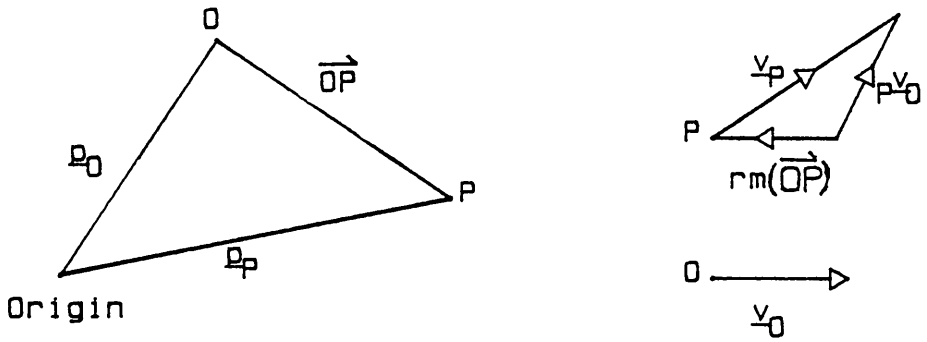


Figure 4.1. The Observation Equations.

The motion resolution problem is the inversion of the equations of observation, i.e. given the relative (left hand side) quantities the absolute quantities on the right must be found. This is clearly underdetermined -- there are two unknowns per equation -- reflecting the ambiguity in the equations: the observation equations are valid in an arbitrary frame of reference and the resolution of the apparent motion into its absolute components depends on the particular frame chosen. The key to motion resolution is the choice of a frame of reference with respect to which "proper motion" will be measured, and in maintaining the stability of that chosen frame over time, despite motion of the observer and the objects.

Underdetermined problems are common in Artificial Intelligence. They arise, for example, in the recovery of shape from shading (Horn and Ikeuchi, 1979), shape from contour (Brady and Grimson, 1981; Barrow and Tenenbaum, 1980; Kanade, 1979), structure from motion (Ullman, 1979), or the computation of optical flow fields using image sequences (e.g. Nagel, 1983). The basic strategy for dealing with underconstrained situations has been to augment the problem, adding constraints until it is determinate or overdetermined. The problem is then solved deterministically. In image understanding

applications, extra constraints are typically obtained from three sources: the physics of the imaging process constrains the local and global properties of images; local indeterminacy may be overcome by global propagation of information (for example, in relaxation methods); criteria which are mathematically, physically, or computationally desirable can be postulated (for example, the smoothness of the velocity field in optical flow extraction) and enforced using the calculus of variations. These methods work to overcome the indeterminacy in the problem.

In the case of motion resolution, there are two profitable extra constraints:

a) a proportion of the visible echo sources are stationary with respect to the world;

b) the proper motions of objects and observer are randomly perturbed smooth motions (the systematic perturbation caused by large scale water currents is fairly smooth).

These two constraints form the basis of the motion resolution schemes presented in this dissertation.

A third, apparently profitable, constraint is:

c) the motions of echo sources are caused by the rigid motion of the underlying objects and are therefore mutually constrained.

However, the relationships between sources and the objects 'carrying' them is unknown -- it is one of the things the interpreter is working to determine -- so this constraint does not help the solution. There are no natural variational constraints to apply and the sparsity of the sonar image makes global propagation of local data difficult.[1]

1. However, it could be argued that global propagation, through space and time, of local relative position and velocity information is just what the motion resolution algorithm described here does.

4.1.2 Exploiting Circularity.

Instead of attempting to overcome the circularity of the motion resolution problem, the solution presented here attempts to exploit it. The basic circularity of the problem is illustrated in Figure

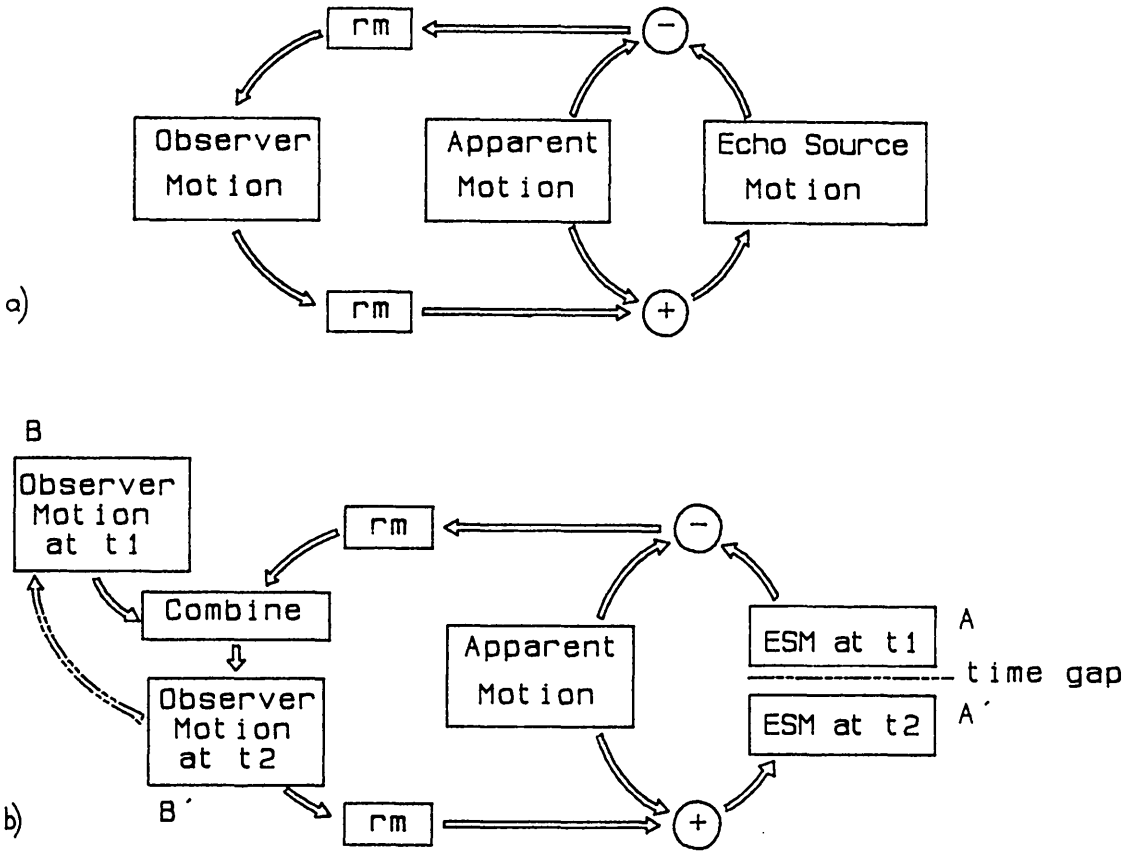


Figure 4.2. Circularity in the Observation Equations.

4.2a. Given the observer motion parameters (position and velocity), the echo source proper motions can be calculated using the reflected motion field (computed from the observer motion) and the apparent motion field (deduced from sensory data). Given the echo source motions, the reflected motion field may be obtained from the apparent motion field and inverted to yield the observer motion parameters.

The way to exploit the circularity is to introduce a time difference into the loop, as depicted in Figure 4.2b. The observations of the apparent motion field are supplied from sensory data in an irregular time series. At each instant there are current estimates of the echo source motion (box A in Figure 4.2b) and observer motion (in box

B) parameters. Using the current estimates of echo source motion, the reflected motion field is calculated from the measured apparent motion field as this becomes available; the reflected motion field is inverted to provide an estimate of the observer motion parameters (in box B'); this new observer estimate defines a new reflected motion field which is combined with the apparent motion field to give new echo source motion estimates (in box A'). The echo source motion estimates in box A are computed from those in box A' of the previous computational cycle, making use of constraint (b) -- the smoothness of echo source trajectories. The computation of new observer motion estimates also takes account of this constraint: values of observer motion computed during previous cycles are incorporated in the current parameter estimation, and the observer motion estimates in box B are predicted from those in box B' of the previous cycle.

The circularity of the problem is thus transformed into a computational cycle which is able to propagate the estimates of observer and echo source motion through time. The steps of the cycle, which I shall call algorithm A, are as follows:

A1 Obtain a measurement of relative position and velocity for an echo source.

A2 Subtract this from the current absolute position and velocity prediction for the source, providing a reflected observer motion estimate.

A3 Combine the reflected observer motion estimates from all simultaneously observed echo sources (inverting the reflected motion field), and use the compound result to update a running estimate of the observer motion.

A4 Add the new estimate of observer motion to the measurements and use the result to update running estimates of the absolute echo source motion.

The running estimates of observer and echo source motion may be constructed using a recursive filter, such as a Kalman tracking filter,[1] which combines previous estimates with those just computed from the new measurements.

4.1.3 Measurement of Relative Position and Velocity.

Motion resolution was defined in Chapter three in terms of the apparent motion of echo sources perceived by the sonar sensor, and equivalently in terms of a viewpoint independent fixed frame of reference. The algorithm for motion resolution described in the previous section was presented in that context. However, an alternative and more uniform definition may be formulated in terms of target tracking.

> Motion resolution is the task of tracking the proper motions of all echo sources, including the observer as a special 'source' that is never seen, with respect to a fixed global frame of reference.

The fixed global (viewpoint independent) frame is established using the tracking data. Implicitly, it is that frame with respect to which the observer is tracked; explicitly, it is the frame in which stationary objects are fixed. Algorithm A may then be seen as a computational mechanism for tracking targets with respect to the viewpoint independent frame, so defining that frame implicitly.

This definition is attractive because, as we saw in Chapter three, target tracking systems are well developed and are able to provide the necessary information for motion resolution processing. By tracking the noisy measurements available from the sonar device, it is possible to estimate the relative position and velocity of each postulated echo source. Confidence information, in the form of error covariances, is provided by the tracking filter for each estimate. The tracking filters are also able to predict the position and

1. The Kalman filter is described informally in Appendix A using an example.

velocity of an echo source at any time and give confidence intervals for their predictions; thus they provide the means of satisfying the design constraints C1, C3 and C4 (section 3.1.1, pages 36ff).

There are two contenders for the tracking filter in a target tracking system: the α-β and the Kalman filters. For motion resolution processing, the intermittent irregular arrival of measurements (constraint C2) eliminates the simple α-β filter (which requires constant intersample time) although the modified filters devised by Holmes (1977) remain a possibility. However, the explicit measurement and state transition models of the Kalman filter (see Appendix A) allow it easily to accommodate fading and Doppler information (if available) on a per-scan basis. The explicit error covariance matrix allows the system to take account of uncertainty in the matching between events and targets. This makes possible a much more robust tracking system.

For linear observer motion the natural choice of reference frame is a Cartesian coordinate frame -- the observer and target motions in this frame correspond to linear state transitions in the tracking filters, so the filters can be made statistically optimal in the sense that their output error covariance is minimal. The Kalman

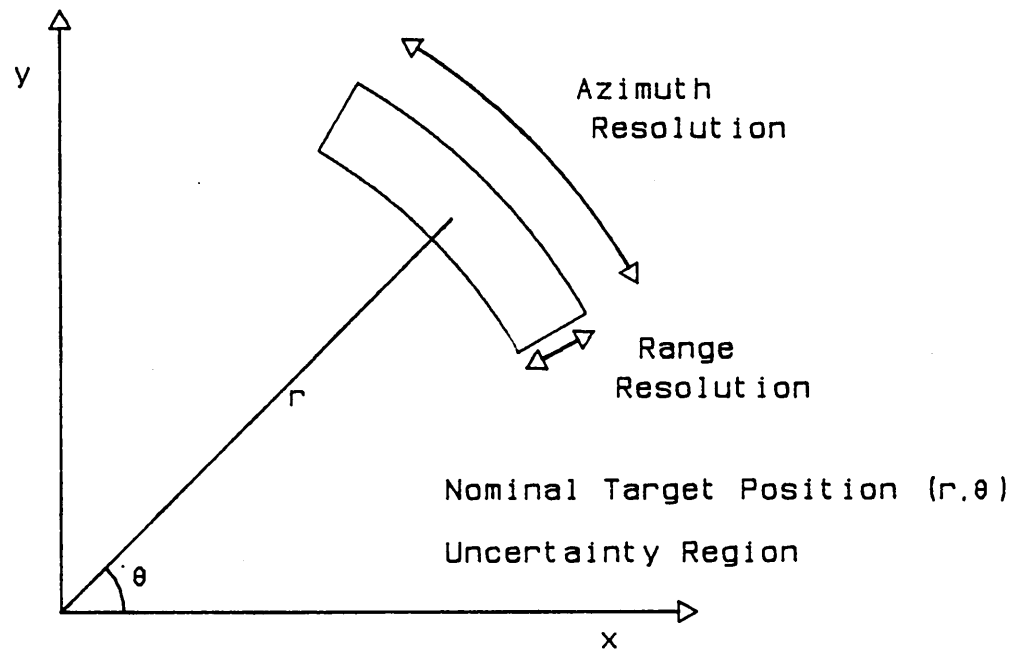


Figure 4.3. Transforming Measurement Uncertainty.

filter accrues further advantages in this Cartesian frame. On transformation from the polar measurement frame to Cartesian coordinates, the constant variance uncorrelated range and bearing noise in the measurements become range dependent errors with bearing dependent correlation between the coordinate dimensions (Figure 4.3). These variable covariance errors rule out the α-β filter, which assumes constant variance uncorrelated measurement noise.

Thus a Kalman filter based tracking system is the natural technique for computing apparent motion estimates (step A1) to drive motion resolution. These filters provide robust optimal noise reduction and are able to keep account of both the correlation between position and velocity estimates and the correlation between the coordinate dimensions.

4.1.4 A Formal Definition of Algorithm A.

This informal description of algorithm A is sufficient to illustrate the structure of the computation. However, for a more precise definition of the computation and a deeper investigation of its properties, it is necessary to formalise the specification of the algorithm. This is done next.

The next three sections define the algorithm in detail and derive certain useful results concerning its stability and expected behaviour. The presentation assumes a fair degree of statistical competence on the part of the reader, and casual readers are advised to read the summary at the end of section 4.1.5 (page 76), where the initialisation procedure for the algorithm is described, and then go directly to section 4.1.7, where the results of the three sections are summarised.

Input.

The first input to the algorithm is a set of echo source relative positions and apparent motions. These quantities are estimated for each source whenever the source is seen, and the input set represents

a number of sources that were sighted simultaneously. Suppose that this happened at time t_n, and denote the input set by $\hat{X}_n$. This set comprises m_n vectors $_i\hat{\underline{x}}_n$ where each $_i\hat{\underline{x}}_n$ consists of relative position ($_i\hat{\underline{p}}_n$) and apparent velocity ($_i\hat{\underline{v}}_n$) estimates for echo source i at time t_n combined as $_i\hat{\underline{x}}_n = [\,_i\hat{\underline{p}}_n \; _i\hat{\underline{v}}_n\,]^T$.

The second input is a corresponding set of estimates of the current observer and echo source positions and velocities in the viewpoint independent reference frame (the absolute frame) at the time t_n when the new input set $\hat{X}_n$ was obtained. Let this set be denoted by $\tilde{\Xi}_n$. It comprises m_n+1 vectors $_i\tilde{\underline{\xi}}_n$ of absolute position and velocity data for the m_n echo sources which contributed to $\hat{X}_n$ (the vectors with $i=1..m_n$) and for the observer (the vector with $i=0$). The tilde on the symbols $_i\tilde{\underline{\xi}}_n$ signifies that they are predictions of the appropriate quantities for time t_n based on data available at times strictly before t_n. The caret, used for example in $_i\hat{\underline{x}}_n$, signifies that the symbol denotes an estimate of the appropriate quantity including information available at time t_n as well as that available strictly before t_n. I shall call this the tilde-caret convention; it is used throughout this dissertation.[1]

Output.

The result of algorithm A is a set $\hat{\Xi}_n$ of observer and echo source absolute position and velocity estimates which incorporate the new information contained in the input set $\hat{X}_n$. $\hat{\Xi}_n$ has m_n+1 members which we denote $_i\hat{\underline{\xi}}_n$, for $i=0..m_n$, following the tilde-caret convention.

1. The tilde-caret convention is also used in much of the literature on statistical filtering. An alternative convention, also used regularly in the literature, expresses the time-dependence explicitly using a conditioning notation. With that convention, an estimated relative state for target i incorporating data obtained at t_n would be $\underline{x}(t_n|t_n,i)$. The parenthesised part of the symbol indicates that the estimate is of the value of the quantity at time t_n (before the "|"), depending on data for target i obtained no later than t_n (after the "|"). As another example, the predicted state of the target based on an estimation at t_m would be $\underline{\xi}(t_n|t_m,i)$.

The Cyclic Computation.

With these input and output definitions completed we now consider the cyclic computation described by steps A2, A3 and A4. In accordance with common mathematical practice I shall simplify the notation by dropping the explicit dependence on time whenever doing so causes no confusion. The implicit dependence on time expressed by the tilde-caret convention remains, however. This shorthand amounts to omitting the post-subscript n. I have also moved the echo source pre-subscript, for aesthetic reasons. Thus m_n becomes m, $\underset{i}{\hat{\underline{x}}}_n$ becomes $\hat{\underline{x}}_i$, $\underset{i}{\tilde{\underline{\xi}}}_n$ becomes $\tilde{\underline{\xi}}_i$ and $\underset{i}{\hat{\underline{\xi}}}_n$ becomes $\hat{\underline{\xi}}_i$.

The individual observer motion estimates constructed by step A2 are just the difference vectors $\tilde{\underline{\xi}}_i - \hat{\underline{x}}_i$ for i=1..m. These vectors comprise an observer position estimate and a reflected motion estimate. Since the observer motion is translational (and without loss of generality $\underline{\omega}_0$ is zero), the latter is just an estimate of the observer's translational velocity. The individual estimates are combined using a weighted summation to yield a composite observer motion estimate, $\hat{\underline{x}}_0$:

$$\hat{\underline{x}}_0 = \sum_{i=1}^{i=m} \beta_i (\tilde{\underline{\xi}}_i - \hat{\underline{x}}_i) \qquad \text{where} \quad \sum_{i=1}^{i=m} \beta_i = 1. \tag{4.2}$$

This combination process accomplishes two things, therefore: it gives an improved observer position estimate; and it inverts the noisy reflected linear motion field. The weight matrices β_i may be chosen to minimise the error covariance of the composite estimate $\hat{\underline{x}}_0$, in which case they are determined by an Information Averaging Filter computation (see Appendix B).

The composite observer motion estimate is combined in step A3 with the observer motion predicted from previous cycles of computation, $\tilde{\underline{\xi}}_0$ using a linear recursive filter (for example, the Kalman filter) to give the observer output estimate:

$$\hat{\underline{\xi}}_0 = \tilde{\underline{\xi}}_0 + \alpha_0 (\hat{\underline{x}}_0 - \tilde{\underline{\xi}}_0). \tag{4.3}$$

The parameter α_0 is the filter gain matrix; for a Kalman filter it is

given by

$$\alpha_0 = \mathbf{cov}\left[\tilde{\underline{\xi}}_0\right]\left\{\mathbf{cov}\left[\tilde{\underline{\xi}}_0\right] + \mathbf{cov}\left[\hat{\underline{x}}_0\right]\right\}^{-1}.$$

Once the new observer motion estimate $\hat{\underline{\xi}}_0$ is available it is used to compute new echo source proper motion and position estimates. The absolute position and velocity estimate for each source is obtained by combining $\tilde{\underline{\xi}}_i$ and $\hat{\underline{\xi}}_0+\hat{\underline{x}}_i$ (which contains the new information obtained at time t_n), again using a recursive linear filter:

$$\hat{\underline{\xi}}_i = (1-\alpha_i)\tilde{\underline{\xi}}_i + \alpha_i(\hat{\underline{\xi}}_0 + \hat{\underline{x}}_i). \qquad (4.4)$$

For a Kalman filter the gain α_i is given by

$$\alpha_i = \mathbf{cov}\left[\tilde{\underline{\xi}}_i\right]\left\{\mathbf{cov}\left[\hat{\underline{\xi}}_i\right] + \mathbf{cov}\left[\hat{\underline{\xi}}_0+\hat{\underline{x}}_i\right]\right\}^{-1},$$

analogously with α_0.

The parameters β_i for i=1..m and α_i for i=0..m, like the inputs and outputs of the algorithm, depend on time -- they are not necessarily constant from cycle to cycle though they are, of course, constant during a particular computational cycle. The choice of these parameters is further discussed in section 4.1.6. Once particular parameters are chosen, equations (4.2), (4.3) and (4.4) define the computation of algorithm A.

Forward Prediction.

The recursive cycle defined by equation (4.2), (4.3) and (4.4) incorporates new information from the relative input set $\hat{X}_n$ into the running estimates of the observer and echo source absolute motions. The input estimates of these absolute quantities, $\tilde{\Xi}_n$, are computed from previous output estimates by a process of forward prediction. This same process is used to compute observer or echo source position and velocity at any time between estimations.

Forward prediction is based on a model of the way the echo sources and observer move. The model is defined by two matrices, a state transition matrix and a state transition noise covariance matrix.

The formal model is as follows. If $_{i}\underline{\xi}_{m}$ is the true state[1] (position and velocity) of an echo source (or the observer) at time t_m, then the state at time t_n, where $t_n \geq t_m$, is given by the equation

$$_{i}\underline{\xi}_{n} = \phi(t_m,t_n)\,_{i}\underline{\xi}_{m} + {}_{i}\underline{q}_{mn}. \tag{4.5}$$

Here $\phi(t_m,t_n)$ is the state transition matrix, describing the (deterministic) evolution of the state between times t_m and t_n, and $_{i}\underline{q}_{mn}$ is a random noise vector, the state transition noise, that describes the stochastic variation of the state between t_m and t_n. The state transition noise covariance matrix, Q_{mn}, is the covariance of $_{i}\underline{q}_{mn}$. Strictly speaking, since sources may have different state transition models, the matrices $\phi(t_m,t_n)$ and Q_{mn} depend on the source; this dependence is neglected here for clarity.

The forward prediction of absolute state vectors uses the state transition matrix and the expected value of the transition noise (it is impossible to predict the actual value of the random vector $_{i}\underline{q}_{mn}$); the transition noise covariance matrix contributes to the covariance of the prediction. Given an estimate of an absolute state $_{i}\underline{\hat{\xi}}_{m}$ at a time t_m, the best linear prediction we can make for the value of the state at t_n is

$$_{i}\underline{\tilde{\xi}}_{n} = \phi(t_m,t_n)\,_{i}\underline{\hat{\xi}}_{m} + E\left[\, {}_{i}\underline{q}_{mn} \,\right]. \tag{4.6}$$

The error covariance of this prediction (the covariance of the difference $_{i}\underline{\tilde{\xi}}_{n} - {}_{i}\underline{\xi}_{n}$) includes a contribution due to the transition noise $_{i}\underline{q}_{mn}$:

$$\mathbf{cov}\left[\, {}_{i}\underline{\tilde{\xi}}_{n} \,\right] = \phi(t_m,t_n)\,\mathbf{cov}\left[\, {}_{i}\underline{\hat{\xi}}_{m} \,\right]\phi(t_m,t_n)^{T} + Q_{mn}. \tag{4.7}$$

Practical forward prediction of states is complicated by two factors. First, sources will be seen for the first time at some time t_n; second, sources are seen intermittently, i.e. there is no guarantee that a given source will be seen at a particular time t_n. Both these factors account for the variation in m_n, the size of the input set $\hat{X}_n$. It follows from this variation that $\tilde{\Xi}_n$ varies in size and

1. Symbols which carry neither caret nor tilde generally denote the true value of a quantity (as opposed to an estimate or prediction of the quantity).

will in general be a different size from the sets $\hat{\Xi}_m$ generated by previous cycles of algorithm A.

This problem is easy to solve. Consideration of new source processing (first time sightings) is deferred to the next section. For sources that have been seen before, the prediction $_{i}\underline{\tilde{\xi}}_n$ at time t_n is computed using equation (4.6) with the estimate $_{i}\underline{\hat{\xi}}_k$ computed at the time t_k when the target was last seen.

Relative Motion Tracking.

The Kalman filters used for target position tracking are defined in a similar manner to those represented by equations (4.3) and (4.4). For each target i the filter estimates a state vector comprising position and velocity components using measurements containing only position components. Suppose that the measurements are $_{i}\underline{\rho}_n$, and that the measurement noise has covariance $_{i}R_n$. The filter state vector is just the vector $_{i}\underline{\hat{x}}_n$ defined above. The new estimated state is computed from a prediction of the state and a new measurement using

$$_{i}\underline{\hat{x}}_n = {}_{i}\underline{\tilde{x}}_n + {}_{i}\gamma_n\left({}_{i}\underline{\rho}_n - {}_{i}H_n\ {}_{i}\underline{\tilde{x}}_n\right). \tag{4.8}$$

The matrix $_{i}\gamma_n$ is the Kalman gain, relating the two component prediction error vector $_{i}\underline{\rho}_n - {}_{i}H_n\ {}_{i}\underline{\tilde{x}}_n$ to its corresponding state correction (recall that the filter operates by predicting the input measurement and then correcting the state estimate appropriately using the error in prediction). The matrix $_{i}H_n$ which relates the predicted state $_{i}\underline{\tilde{x}}_n$ to the predicted measurement $_{i}H_n\ {}_{i}\underline{\tilde{x}}_n$ is part of the explicit measurement model (the other part of the model is $_{i}R_n$, the measurement error covariance matrix). The Kalman gain $_{i}\gamma_n$ is calculated, analogously with α_0 in equation (4.3), by

$$\begin{aligned} {}_{i}\gamma_n &= \mathbf{cov}\left[\ {}_{i}\underline{\tilde{x}}_n\ \right] {}_{i}H_n^T \left\{\mathbf{cov}\left[\ {}_{i}\underline{\rho}_n\ \right] + \mathbf{cov}\left[\ {}_{i}H_n\ {}_{i}\underline{\tilde{x}}_n\ \right]\right\}^{-1} \\ &= {}_{i}\tilde{\pi}_n\ {}_{i}H_n^T \left\{{}_{i}H_n\ {}_{i}\tilde{\pi}_n\ {}_{i}H_n^T + {}_{i}R_n\right\}^{-1}, \end{aligned} \tag{4.9}$$

where

$$ {}_i\tilde{\pi}_n = \mathbf{cov}\left[{}_i\tilde{\underline{x}}_n \right]. $$

The error covariance of the estimate ${}_i\hat{\underline{x}}_n$ is then given by

$$ {}_i\hat{\pi}_n = \left(1 - {}_i\gamma_n \, {}_iH_n\right) {}_i\tilde{\pi}_n. \qquad (4.10) $$

Finally, the state predictions and its error covariance for time t_n are computed from the state estimate and its error covariance at time t_m using equations analogous to (4.6) and (4.7).

4.1.5 Initialising the Cyclic Computation.

The computation defined by equations (4.2) to (4.6) is recursive: it computes current estimates of observer and echo source proper motion given a new set of relative motion measurements ($\hat{\underline{X}}_n$) and previous values of the observer and echo source absolute state estimates (from which the values in $\tilde{\Xi}_n$ are computed by forward prediction). To complete the definition of algorithm A it is necessary to specify suitable initial conditions for the recursion.

The initialisation is motivated by the following pair of theorems which may be proved independently of the actual values of the gain matrices α_i and weight matrices β_i used in the formalisation given above. (Proofs of these theorems are given in Appendix C.)

Theorem 4.1. The Bias Propagation Theorem.

> Suppose that the observations of relative motion $\hat{\underline{X}}_n$ obtained at time t_n are unbiassed[1] but that the absolute state predictions in $\tilde{\Xi}_n$ all contain a (common) bias $\underline{b}_n$. Then the absolute state estimates in $\hat{\Xi}_n$ also contain bias $\underline{b}_n$.

1. An estimator is said to be biassed if its expected value differs from the expected value of the quantity it is estimating. Thus for an unbiassed estimator, the estimation error has zero mean.

Theorem 4.2. The Bias Prediction Theorem.

Suppose that an estimated absolute state at time t_m contains a bias $\underline{b}_m$. Then the corresponding predicted absolute state for time t_n contains a bias $\underline{b}_n$ where

$$\underline{b}_n = \phi(t_m,t_n)\underline{b}_m. \qquad (4.11)$$

It should be noted that the absolute state referred to in Theorem 4.2 may be an observer or echo source absolute state, and that the bias vectors $\underline{b}_m$ and $\underline{b}_n$ in both theorems include both position and velocity components. Strictly speaking, the value of the state transition matrix, $\phi(t_m,t_n)$, used in equation (4.11) must be the appropriate one for the particular target concerned.

Theorem 4.1 confirms the appropriateness of a natural procedure for initialising the absolute state estimates of new sources using an existing observer absolute state estimate. The procedure is this: for each source first detected at time t_n, omit the source from the recursive computation and instead set the output estimate $_i\hat{\underline{\xi}}_n$ for that source using

$$_i\hat{\underline{\xi}}_n = {}_0\hat{\underline{\xi}}_n + {}_i\hat{\underline{x}}_n. \qquad (4.12)$$

Note that this is a degenerate case of equation (4.4) with the gain α_i set to unity. This initialisation procedure is attractive because, given the initial conditions of Theorem 4.1 at time t_n, the conclusion of that theorem also holds for the newly initialised state estimate; the bias present in the observer absolute state $_0\hat{\underline{\xi}}_n$ is introduced directly into $_i\hat{\underline{\xi}}_n$ by equation (4.12).

When this initialisation procedure is used an important result follows from the two theorems given above. This is the initialisation theorem given below (also proved in Appendix C).

Theorem 4.3. The Initialisation Theorem.

Suppose that an estimate of the observer absolute state is available for some time t_0 and that it contains bias $\underline{b}_0$. Suppose further that no echo sources are known at t_0

and that the state transition matrix depends only on the time (and not on the target). Then, using the initialisation procedure (4.12), the estimated or predicted absolute state vectors for all echo sources (and the observer) known at time $t_n \geq t_0$ contain a common bias $\underline{b}_n$ given by

$$\underline{b}_n = \phi(t_0, t_n)\underline{b}_0. \qquad (4.13)$$

Theorem 4.3 shows, informally, that the stability of the viewpoint independent reference frame (which is implicitly defined by the observer absolute state values) depends only on the initial bias in the observer absolute state. If the initial absolute frame origin has a motion defined by the bias vector $\underline{b}_0$ at time t_0 it continues to move in accordance with that initial motion. The recursive computation of algorithm A therefore maintains the initial motion of the viewpoint independent reference frame as it propagates and improves the absolute state estimates for the observer and echo sources.

There are two points to note with regard to Theorem 4.3. First, the theorem does not imply that perfect initialisation of the observer absolute state will ensure that the absolute frame remains stationary at all times. The input measurements and the absolute state predictions and estimates all contain noise components which cause the viewpoint independent frame to move about. Its origin position is subject to perturbation by noise. The theorem is not about actual instantaneous motion; rather, it is about expected motion. Thus the average motion of the absolute reference frame is determined by the observer initialisation, but the precise detail of its instantaneous motion is random.

The second point is similar to the first. The bias motions in terms of which Theorem 4.3 is expressed are also expected motions. Thus the stability of the absolute frame does not depend on the perfect initialisation of the observer absolute state vector, but rather on its unbiassed initialisation. The initialisation may contain noise whose expected value is zero.

The initialisation problem has been reduced, by Theorem 4.3, to the problem of finding a single unbiassed estimate of the observer position and velocity at some instant of time t_0.

Unless the viewpoint independent reference frame is required to register with that input in some a priori map, its origin is arbitrary. Even if a map is available, a reasonable strategy is to assume an arbitrary initial origin and process sensory data for a time, then attempt to fit what has been seen to the map, shifting the origin appropriately if a good fit can be found. Given that the viewpoint independent frame origin is arbitrary a suitable initialisation procedure for it is to define the current observer position at some time t_0 (that position is of course unknown) to be the origin. A natural instant to chose is a moment when a set of targets is observed, since the absolute positions of these targets can then be set from their relative position measurements, using equation (4.12).

A suitable observer velocity initialisation is harder to find. It is also more critical than the position initialisation, since a bias in the observer velocity estimate will cause the viewpoint independent reference frame to drift.

If a set of echo sources with known proper motions were available, the observer velocity could be determined by inverting the reflected motion field using (4.2). However, there is no direct way of measuring the absolute velocity of an echo source so the reflected motion field cannot be computed. The way forward is provided by the constraint that a proportion of the visible echo sources at any time are stationary in the world (constraint (a) above). These stationary echo sources can be used to initialise the observer velocity.

Unfortunately, the interpreter cannot identify the stationary echo sources directly. A source is only known to be stationary if it is fixed in the viewpoint independent reference frame, which is still to be defined. The system must therefore discover, by indirect means, which of the sources observed are stationary.

Knowing that a proportion of the sources seen at a given time are stationary, the motion resolution algorithm may assume, as a first approximation, that all sources seen are stationary. With this assumption a reflected motion field can be deduced from the apparent motion of the sources. This reflected motion field may be inverted using an Information Averaging Filter to extract the common observer motion from the noisy reflected motion values[1] (equation (4.2)) and the observer initial velocity may be set using this value (by application of equation (4.3) with α_0 equal to unity).

This initialisation procedure is satisfactory provided that the echo sources first seen are stationary. If any of those sources are moving, the initial observer velocity estimate is biassed by an amount equal to the mean proper velocity of the set of sources. This bias arises because the errors in the reflected motion estimates for the moving sources are no longer zero mean: they contain a bias equal to the source absolute velocity.

Since it is unlikely that the initial source set will contain only stationary sources the above procedure is unsatisfactory as it stands. It can be improved by attempting to eliminate moving sources. A statistical hypothesis test, which attempts to detect bias in the reflected motion errors for an individual source, provides the necessary information. Sources are eliminated from consideration only when they have a statistically significant error velocity under the stationary source hypothesis. The confidence information required for testing the hypothesis is available from error covariance information propagated by algorithm A (the details of this are given in the next section).

A statistical approach of this type raises two new difficulties. First, the position and velocity uncertainties associated with the source data in the initial set of sources will be large because there is little information available. The hypothesis test is therefore unlikely to discover any significant bias in the reflected motion

1. The noise, derived linearly from zero mean noise in the input event positions provided by the Sonar Device Interface, has zero mean so the expected value of the reflected motion field will be exactly invertible, provided that the sources are actually stationary.

field. The same test done later, when more information has been collected, would be more accurate and reliable.

Second, a hypothesis test done once at initialisation time is obviously misleading because a source may change its state and begin to move, after the test has decided that it is stationary. The solution to this problem is to allow the status (stationary or moving) of each source to vary depending on the conclusion of a hypothesis test done after each computation of the state vector, $_{i}\hat{\underline{\xi}}_{n}$, for that source, and to distribute the initialisation of the recursive computation over several cycles of the algorithm.

Unfortunately, the algorithm as it stands cannot be initialised in a distributed manner because of the consequences of Theorem 4.3; under that theorem only the first initialisation cycle is effective. The remedy for this difficulty is to alter the algorithm to take explicit account of the source status. For moving sources, no changes are necessary. For sources that are stationary (according to the hypothesis test, of course) the estimated absolute velocity components in the output state vector $_{i}\hat{\underline{\xi}}_{n}$ for those sources are set artificially to zero after using equation (4.4). This action uses the fact that the true stationary sources have no absolute velocity.

With this extension Theorems 4.1 to 4.3 no longer apply in general. However, if the absolute state estimates are unbiassed at time t_m and no targets are misclassified as stationary during the interval $[t_m, t_n]$, then the state estimates are also unbiassed at time t_n. This follows from Theorem 4.3, which does hold under these conditions. If moving targets are misclassified as stationary the operation of the recursive cycle actually introduces bias into the absolute state estimates.

An interval during which sources are misclassified as stationary will be termed an initialisation phase. An example of such a phase occurs during the initial period of operation of the algorithm where all sources are assumed stationary and moving sources have yet to be identified. During an initialisation phase the origin of the viewpoint independent reference frame will drift in a manner depending on source and observer motion. Once the moving sources are

rejected by the hypothesis test, and are correctly classified, the reference frame will be "locked" to the stationary sources by the recursive cycle. Since the reference frame origin is arbitrary this drift does not matter during the initial period of operation (if the algorithm re-enters an initialisation phase at some later time, the drift is a serious problem).

Summary.

In summary, two procedures are used for initialising the observer and echo source absolute states. First, for those sources seen when the observer absolute state estimate is available, the source absolute state is initialised by setting its absolute state estimate to the sum of the current observer absolute state and the initial measurement vector (i.e. using equation (4.12)). The natural initial status for such sources is 'moving'. If they are actually stationary sources, the hypothesis test will detect that; if not, they are correctly classified and therefore do not force an erroneous initialisation phase into the operation of the algorithm. An initialisation phase occurs when a moving target is misclassified as stationary, and during such a phase the viewpoint independent reference frame origin is not stable.

Second, the absolute states of the observer and first set of sources seen are initialised by setting the initial status of all the sources to 'stationary' and by setting the absolute source position components equal to the corresponding relative position components. The observer state estimates are then initialised by setting the observer absolute position to zero and its absolute velocity to a weighted composite of the measured relative velocities of the targets that have just been seen (i.e. by using equation (4.2), followed by equation (4.3) with α_0 set to unity). If all the initial sources were actually stationary this initialisation leaves the algorithm operating correctly; if not, the algorithm remains in the initialisation phase (because there are moving objects classified as 'stationary') until the moving sources are correctly classified.

4.1.6 Choosing the Gain and Weight Parameters.

All that remains to complete the description of algorithm A is a mechanism for setting the values of the gain and weight matrices required by equations (4.2), (4.3) and (4.4).

In fact, the problem of determining the gains, α_i, is just the problem of choosing suitable weights (β_i) for combining a pair of quantities. This is seen most clearly by comparing equations (4.2) and (4.4) -- the latter combines a predicted absolute state $\underline{\xi}_i$ (weight $1-\alpha_i$) with a newly computed absolute state $\hat{\underline{\xi}}_0+\hat{\underline{x}}_i$ (weight α_i). In what follows, the terms 'gain' and 'weight' are used interchangeably.

The basic rationale underlying the choice of gain matrices is that those estimates with large error covariance should contribute less to the result than those with smaller error covariance. Ideally, the weighted composite should have as small an error covariance as possible given the error covariances of the estimates being combined.

Given that the individual estimates contain uncorrelated zero mean errors, the optimal composite $\overline{\xi}$ is defined by the Information Averaging Filter equation (defined in Appendix B)

$$\overline{\xi} = \overline{P} \sum_{i=1}^{i=n} P_i^{-1} \xi_i, \quad \text{and} \quad \overline{P} = \left\{ \sum_{i=1}^{i=n} P_i^{-1} \right\}^{-1}, \tag{4.14}$$

where P_i is the error covariance of the estimate ξ_i. $\overline{P}$ is the error covariance of the composite estimate.

There are two difficulties in trying to choose optimal gains and weights for algorithm A. First, optimal gain calculations require an accurate knowledge of the error covariances of all the quantities being combined. The maintenance of exact covariance estimates requires a complex computation which is undesirable in relation to the strong real-time constraint on sonar interpretation. Second, the optimal gain computation is further complicated by the fact that the various estimates to be combined are not statistically independent. The noises in the observer motion estimates $\underline{\xi}_i - \hat{\underline{x}}_i$, for example, are

correlated because, in general, the $\underline{\tilde{\xi}}_i$ are based on common measurements at some point in their past.

In view of the prohibitive complexity of optimal gain calculations, two approximations can be made. First, upper bounds to the covariances of terms such as $\underline{\tilde{\xi}}_i - \underline{\hat{x}}_i$ are used. This saves the need to account for the correlation between the parts of the term.[1] Second, correlation between individual estimates is neglected and composite estimates are computed using the relatively cheap Information Averaging Filter, as given by equation (4.14).

Given these approximations, the weight and gain matrices are computed as follows. Let $\tilde{P}_i$ and $\hat{P}_i$ be the estimates of the error covariances for the state vectors $\underline{\tilde{\xi}}_i$ and $\underline{\hat{\xi}}_i$ for i=0..m. Suppose further that the error covariance of the input estimate $\underline{\hat{x}}_i$ is $\hat{\pi}_i$, for i=1..m. The upper bound used for the error covariance of the individual observer estimate $\underline{\tilde{\xi}}_i - \underline{\hat{x}}_i$ is $2(\tilde{P}_i + \hat{\pi}_i)$ (see Appendix C) and therefore

$$\beta_i = 2\hat{\pi}_0(\tilde{P}_i + \hat{\pi}_i)^{-1}, \quad \hat{\pi}_0 = 2\{\sum_{i=1}^{i=m}(\tilde{P}_i + \hat{\pi}_i)^{-1}\}^{-1}, \tag{4.15}$$

where $\hat{\pi}_0$ is the error covariance computed for $\underline{\hat{x}}_0$.

The gain α_0 is computed using the Kalman gain formula. The input covariance is $\hat{\pi}_0$ and the predicted state covariance is $\tilde{P}_0$; the Kalman formula gives:

$$\alpha_0 = \tilde{P}_0(\tilde{P}_0 + \hat{\pi}_0)^{-1}, \quad \hat{P}_0 = (1-\alpha_0)\tilde{P}_0. \tag{4.16}$$

The gains, α_i, for equation (4.4) are computed analogously; $\tilde{P}_i$ replaces $\tilde{P}_0$, the estimated error covariance of $\underline{\hat{\xi}}_i$ is $\hat{P}_i$ analogously to $\hat{P}_0$, and the input covariance is the upper bound $2(\hat{P}_0 + \hat{\pi}_i)$ rather than $\hat{\pi}_0$. Forward prediction of the covariances, calculating $\tilde{P}_i$ from previous $\hat{P}_i$ for i=0..m, proceeds as indicated by equation (4.7).

1. In the case given, $\underline{\tilde{\xi}}_i$ contains earlier measurements from the sequences of which $\underline{\hat{x}}_i$ is a member, and the members of this sequence (generated by a Kalman filter) are correlated.

4.1.7 Summary.

At this point it is appropriate to pause and take stock of the situation. A motion resolution algorithm has been described informally and defined formally as a recursive linear filter. This is algorithm A. It defines and maintains a stationary viewpoint independent reference frame by tracking the apparent motions of echo sources, deducing their absolute (viewpoint independent) motions, and deducing the observer absolute state position and velocity, which implicitly define the absolute frame origin.

The formal definition of the algorithm, presented in sections 4.1.4 to 4.1.6, specifies how to:

- construct relative position and apparent motion estimates from the position measurements supplied by the sonar device;
- initialise the recursive computation of algorithm A;
- implement the recursive computation in terms of Kalman tracking filters and Information Averaging Filters;
- approximate the error covariance matrices of state estimates in a computationally cheap manner.

It should be noted that few assumptions regarding the state transition model ($\phi(t_m,t_n)$, $_i\underline{q}_{mn}$, Q_{mn}) are implicit in the formal definition of algorithm A. For the purpose of description it was tacitly assumed that the state vectors comprised position and velocity components. In fact, the recursive filter does not require state vectors of that form; the only constraints on the dimension and content of state vectors are that the resulting algorithm be dimensionally consistent and physically sensible (although this latter is not, strictly speaking, necessary!). The precise nature of $\phi(t_m,t_n)$ and $_i\underline{q}_{mn}$ are similarly unconstrained, except that if a Kalman filter is used as the realisation of step A3 or A4 then $\phi(t_m,t_n)$ must be

independent of the state vector value and $\underline{i}_{mn}$ should be drawn from a unimodal zero mean distribution. The input measurement sequence $\{\underline{z}_n\}$ (or its equivalent in terms of position-only measurements) is also flexible; its definition was cast as a sequence of sets arriving at a sequence of times. Therefore it can comprise irregular intermittent data if desired.

So far it has been shown that in some circumstances the estimates of absolute states behave consistently. However, a number of things have <u>not</u> been demonstrated formally.

- Nothing has been said about the noise reduction properties (if any) of the algorithm.

- No attempt has been made to prove rigorously that the algorithm is stable (indeed, such a proof is really outside the scope of this thesis).

- It is not clear how successful the initialisation strategy will be in practice.

These points can be addressed informally by a qualitative consideration of the recursive loop equations (4.2), (4.3) and (4.4). The presence of true stationary targets classified correctly will force bias towards zero in the absolute state vectors. Moving targets classified correctly (and stationary targets classified incorrectly) will, in the absence of any correctly classified stationary targets, behave as indicated by Theorem 4.3 (irregularities in the bias vector distribution will tend to be smoothed out, since the weights and gains are less than unity in some sense). The tentative conclusion is that the system will be stable and some noise reduction may be expected provided that no moving targets are misclassified after the startup initialisation phase and provided that some stationary targets are correctly classified.

4.2 Preliminary Experiments with Algorithm A.

The points raised in section 4.1.7 were addressed empirically in two stages. First, a testbed implementation of the algorithm was constructed and some preliminary experiments were done. These are described below. Subsequently, the algorithm was reimplemented taking advantage of the insight gained through these experiments (section 4.3 describes this new implementation) and a number of other experiments, including Monte Carlo tests, were done to assess the new system's performance. These are described in section 4.4.

4.2.1 The Experimental World Model.

All the performance experiments described in this thesis use simulated data obtained from a simple, yet realistic, world model. There are three major reasons for this use of simulated rather than real sonar data.

First, the simulated data used is two-dimensional. It is also possible to restrict the class of target motions permitted in the simulation. Real sonar data, collected by the ANGUS vehicle for example, is intrinsically three-dimensional and the targets and vehicle can make arbitrary three-dimensional motions including both translational and rotary observer movement. Using this type of input, the general motion resolution problem would have to be solved in one step.

Second, data generated by simulation is repeatable and precisely controllable. The motion and visibility parameters of targets and the measurement noise characteristics can be varied in a controlled manner and the 'true' positions, velocities, etc., of the observer and targets are available for comparison with the outputs of the motion resolution system.

The third, related, reason is that performance testing of complex statistical algorithms such as algorithm A requires a large bulk of controlled data. The performance is assessed by 'sampling': a number

of situations are chosen at random and the algorithm is tested in each case. The performance for the random sample is then taken as indicative of the performance of the algorithm. The accurately controllable output of a simulation is ideally suited to such tests.

The world model chosen for simulation was a plane containing a number of independent point targets, one of which is designated the observer. The targets represent echo sources which, for the purposes of motion resolution, are independent points. Each target had two degrees of freedom, parameterised by its position coordinates with respect to the world origin (in Chapter five, where observer rotation is considered, the observer has an extra rotary degree of freedom). To model fading and the limited resolution of sonar equipment targets were intermittently visible and measurement noise from a Gaussian distribution was added to the relative range and azimuth of each currently visible target before the data was passed to the motion resolution system. The probability of appearance of a target was constant, set for each target, and visibility was determined using a Bernouilli random variable (one that takes values zero or one).

Target motions were modelled by linear constant acceleration trajectories. Random velocity and acceleration components could be generated if required, modelling the random perturbation of target motion by small-scale currents. The effect of large scale water currents or deliberate target motions is adequately modelled by the linear deterministic component of the trajectory.

The simulation parameters -- target visibility and trajectory parameters and the measurement noise variances -- are defined at initialisation time, but can be changed at any time during the simulation. The output of the simulation is a sequence of events; the true measurements and target position and velocity data are also available at any time.

4.2.2 The Experimental Implementations of Algorithm A.

In this preliminary experiment, four alternative realisations of algorithm A were tested. The implementations differed in the actual computation done during the recursive loop. The degree of success of an implementation was assessed directly by comparing, each time a target was seen, the actual motion parameters computed by the simulator with the corresponding components of the estimated absolute state vector for that target. For each quantity compared, the extremal values were noted and the sums of the errors and of the square errors were accumulated. The criteria for evaluation were:

a) accuracy -- the degree of consistency between the computed motion parameters and the world motion parameters;

b) noise reduction -- the ratio of the covariances of computed quantities to raw input variances, and the relationship between actual error covariances and the covariances predicted by the motion resolution algorithm;

c) robustness -- the behaviour of the algorithm when faced with increasingly dominant noise or with increasingly complex world behaviour.

All four implementations used the same state transition model. The state vectors comprised position, velocity and acceleration components for a single dimension. The correlation between the spatial dimensions was neglected, so each dimension was filtered independently. Target motions were taken to be constant acceleration linear trajectories; the random motion component of the trajectories was not modelled for the Kalman filters in these preliminary tests (i.e. the state noise covariance matrix was taken to be zero).

The implementations also allow a total of ten observations of a target, during which time the start-up transients in the relative tracking filter die away, before it is included in the cyclic motion

resolution scheme. The tenth observation of a target therefore prompts the initialisation of the target absolute state vector, and, if necessary, the initialisation of the observer absolute state also. For these tests the segmentation problem (matching events to the targets the caused them) was ignored -- the target causing each event was recorded by the simulator.

Exploratory tests using various viewing probabilities showed that the accuracy and noise reduction properties of the algorithm were affected strongly by the total number of observations obtained for a target but only weakly by their rate of acquisition. This result was not surprising: the noise performance depends principally on the Kalman filter gains, which are independent of the actual values of measurements but depend on the number of times the filters have processed input measurements. With full segmentation processing, where the filter gains depend indirectly (via the matching decisions taken by segmentation) on the measurements obtained and their times of arrival, this result will no longer hold. In view of the small effect of the target data rate, the viewing probabilities were set to unity in subsequent tests to maximise the data throughput for error logging.

It was also found that ten observations were sufficient to stabilise the relative motion estimates for targets, and that a pessimistic value of the measurement noise variance was necessary in the Kalman filter measurement model to ensure the filter's numerical stability. The actual value used in the tests was five times the true measurement noise variance; the implications of this result will be discussed later.

Implementation 1.

The first implementation followed the description of algorithm A given above with two modifications: the absolute state estimation filters were passed position and velocity estimates only (the acceleration components were tracked but not propagated round the recursive loop); and the Information Averaging Filter used to combine

observer motion estimates was uncoupled, i.e. the correlation between position and velocity noises was neglected. The former difference changes slightly the Kalman filter equations in (4.3) and (4.4) -- they must now include an explicit measurement matrix (like equation (4.8)). The latter change amounts to making the weights β_i in (4.2) diagonal matrices.

The algorithm was found to converge reliably and remain stable for at least 1000 seconds of simulated time. The filter estimates of the error covariances were uniformly larger than the true values logged, though this was not surprising in view of the pessimistic measurement noise variance used in the filter measurement model.

The errors in the target absolute states were larger than those in the inputs to the absolute state filters. This was a surprising result, as it suggested that the absolute state tracking filters were aggravating the noise present in the state estimates rather than reducing it. Implementation 2 attempted to resolve this problem.

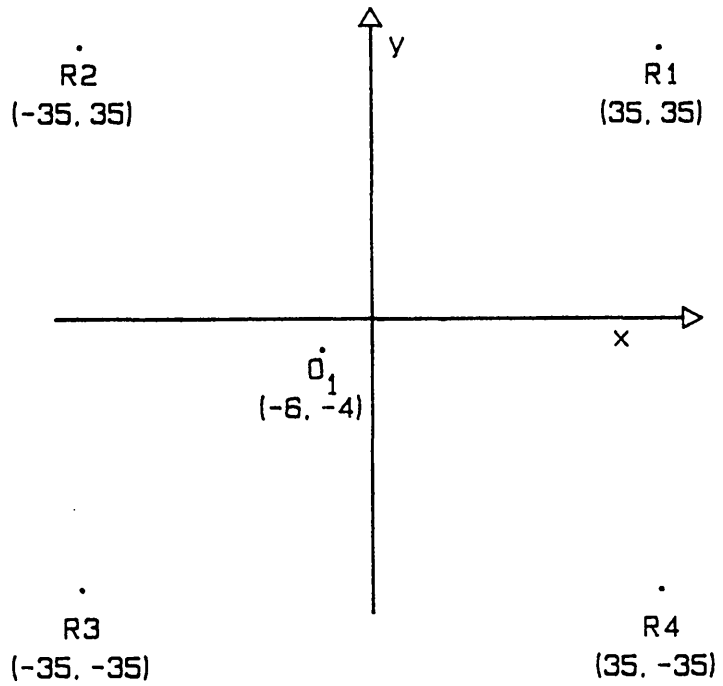


Figure 4.4. The Initial Target Layout for a Test Run.

A typical set of results obtained from implementation 1 are given in Tables 4.1 and 4.2. The initial arrangement of targets and observer for the test, which ran for 100 seconds of simulated time,

are shown in Figure 4.4. Targets one and four were stationary while targets two and three were moving with velocity (0.35,0.2) and (0,0.6) respectively.

The measurement noise parameters chosen were of a medium quality sonar:[1] the range resolution was ±3 centimetres and the angular beam width about 3.4°. The sonar repetition rate was 2.5 Hz.

	Measurement	Absolute State Tracking Filter		
Target	Position/m^2	Pos. In/m^2	Pos. Out/m^2	Vel. Out/m^2s^{-2}
Observer	-	-	0.126	1.94e-3
Target 1	0.136	0.129	0.147	2.33e-3
Target 2	0.177	0.183	0.259	4.00e-3
Target 3	0.0214	0.141	0.167	7.63e-3
Target 4	0.118	0.125	0.147	4.07e-3

Table 4.1. X Component Error Variances for Implementation 1.

Table 4.1 summarises the error variance values for the x components of the measurement absolute state input, and absolute position and velocity estimate vectors for the observer and targets in the test. The increase in noise variance across the absolute state tracking filter is clearly visible.

	Position Estimate Offset/m				
	Observer	Target 1	Target 2	Target 3	Target 4
x	1.55	1.63	1.84	1.66	1.66
y	2.60	2.73	2.78	2.75	2.69

Table 4.2. Mean Position Estimate Offsets.

Table 4.2 gives the values of the mean error in absolute position for each target in the test. Note that the value of the observer mean error is arbitrary (it depends on the drift in the viewpoint independent reference frame origin which occurs during the initialisation phase and the initial observer coordinates given to the simulator) but the means for the targets should be consistent with the

1. An example of a comparable commercial sonar is the SEAVISION system (Duck, Goodman and Griffiths, 1984) which has a range resolution of 20 to 30 cm and a beamwidth of 1.5 to 2 degrees.

observer mean. The degree of inconsistency in these values reflects the spatial inconsistency of the absolute reference frame.

Implementation 2.

This implementation was prompted by the observed increase in error variance across the absolute tracking filter. Implementation 2 was identical to implementation 1 in every respect but one: the gains α_i for the absolute tracking filters were set to unity in equation (4.4), effectively eliminating the tracking action of that filter.

Implementation 2, like the first implementation, converged reliably to a stable set of absolute state estimates. It performed better, in terms of accuracy and noise reduction, than did its predecessor. The improvement appeared to be the result of the faster response time of the loop without the absolute tracking filter.

The results obtained from implementation 2 for the test described as an example under implementation 1 are given below in Tables 4.3 and 4.4. Comparison of these tables with Tables 4.1 and 4.2 shows the change in performance.

	Measurement	Absolute State	
Target	Position/m^2	Position/m^2	Velocity/m^2s^{-2}
Observer	-	0.109	1.82e-3
Target 1	0.136	0.109	1.75e-3
Target 2	0.177	0.128	3.95e-3
Target 3	0.0214	0.116	5.08e-3
Target 4	0.118	0.107	2.56e-3

Table 4.3. X Component Error Variances for Implementation 2.

Table 4.3 presents the x component error variances for implementation 2. Note that there is now no separate absolute state position input column (since the gain is one the filter output is now equal to its input). The test exhibits noise reduction for all targets except number three which has an unusually low x variance (its y variance is 0.116) and the absolute state estimate noise variances are uniformly smaller than the equivalent entries in Table 4.1.

	Position Estimate Offset/m				
	Observer	Target 1	Target 2	Target 3	Target 4
x	1.29	1.25	1.38	1.29	1.29
y	1.93	1.95	1.95	1.94	1.92

Table 4.4. Position Estimate Offsets for Implementation 2.

Table 4.4 gives the position estimate offset values for the test under implementation 2. There is good agreement between the values, the only anomaly being the x offset of target two, showing good spatial consistency of the absolute reference frame. A comparison of these figures with the corresponding entries in Table 4.2 shows the extent of the improvement in spatial consistency (recall that only the agreement among the values, rather than the actual numbers, is significant in assessing spatial consistency).

Implementations 3 and 4.

These test implementations were identical with implementations 1 and 2 respectively except that the Information Averaging Filter (equation (4.2)) was coupled, allowing for the correlation between position and velocity noises. In terms of the weights β_i, the restriction that they be diagonal was relaxed in these implementations.

Neither of these versions offered a significant improvement over their equivalent uncoupled implementations.

Further Results for Implementation 2.

The main conclusion to be drawn from these experiments was that implementation 2 performed satisfactorily and was significantly better than the original version of the algorithm. Further testing of this modified algorithm (α_i set to one) was therefore carried out.

The test world comprised four targets, initially arranged at the corners of a 70 metre square centred on the origin (this arrangement was inspired by the arrangement of legs on an imaginary quadrupedal

oil platform), and the observer which was initially near the origin. The sonar repetition rate and range resolution were 2.5 Hz and ±3 centimetres, and the beam width was 30° in the first test (a very poor sonar!) and 3° in the latter three tests reported below.

Variances (Max/Min)

Test	Measurement Position/m^2		Absolute State Position/m^2	Velocity/m^2s^{-2}
1	12/10	T	1.8/0.5	0.2/0.04
		O	0.45/0.4	0.11/0.08
2	0.177/0.020	T	0.128/0.040	7.3e-3/1.8e-3
		O	0.110/0.040	7.1e-3/1.8e-3
3	0.450/0.002	T	0.120/0.012	3.5e-3/1.0e-3
		O	3/1.8	1.6e-4/1.3e-4
4	0.180/0.140	T	0.156/0.054	7.3e-4/3.2e-6
		O	0.600/0.210	3.2e-5/2.7e-5

T = Target O = Observer

Table 4.5. Variance Summary for Tests 1 to 4.

Table 4.5 gives a summary of the variance information for Tests 1 to 4. The details of the tests were as follows.

Test 1. The four targets were stationary, the observer velocity was (0.06,0.08), and the sonar beamwidth was 30°. The test ran for 100 seconds.

Test 2. This was just the test illustrated above by Figure 4.4, with conditions exactly as described under implementation 1, i.e. two moving targets and two stationary ones.

Test 3. The four targets were stationary and the sonar beamwidth was 3°. The test ran for 200 seconds and the observer trajectory was described by a linear velocity of (0.06,0.08) to which a random position perturbation of 0.04 metres standard deviation (for a 0.4 second sample time) was added to each coordinate.

Test 4. This was identical to test 3, but ran for 1000 seconds.

For each test Table 4.5 gives the maximum and minimum variances of the errors in the event positions (relative to the observer), the target absolute position and velocity and the observer absolute position and velocity. The variances are based on a sample of some 240 measurements (2.5 Hz times 100 seconds less the 10 initialisation events). Some noise reduction can be seen in test 1 and 2; tests 3 and 4 are harder to assess because of the random observer movement, but the velocity estimates formed by the algorithm have low variances and the drop in position variance between test 3 and test 4 suggests that much of the larger variance for the observer position in test 3 is due to the initialisation phase. The targets show some noise reduction in both these tests.

	Observer		Target 1		Target 2		Target 3		Target 4	
Test	x	y	x	y	x	y	x	y	x	y
1	3.3	-0.6	2.8	-0.1	4.3	-0.2	2.7	-1.1	4.1	-0.3
2	1.2	1.9	1.1	1.9	1.3	1.9	1.3	2.0	1.2	2.0
3	30.0	39.8	29.9	39.9	30.0	39.8	30.0	39.7	30.0	39.9
4	30.6	40.2	30.5	40.1	30.6	39.9	30.3	39.9	30.6	40.0

Table 4.6. Mean Position Offsets (cm) for Tests 1 to 4.

Table 4.6 gives the position offsets obtained for tests 1 to 4. The stability of the reference frame in each test is indicated, as before, by the similarity between corresponding displacements for different targets and observer. Test 4 is a longer run of test 3, so comparison between these rows shows the time stability of the reference frame; comparison within the rows shows the spatial stability of the frame.

4.2.3 Discussion of Results.

Of the four experimental implementations, number 2 performed the best. It was able to reject target motion, maintain a stable reference frame under extreme error conditions, and under moderate conditions it provided a temporally stable and spatially consistent

reference frame based on only four targets. Thus implementation 2 satisfies the requirements of motion resolution in the situations tested so far, and this result confirms the validity of the motion resolution algorithm.

The most interesting result arising from this test is that algorithm A's performance is significantly improved by modifying equation (4.4) so that the gains α_i are unity for all targets. This variation of step A4, stated explicitly below, will be denoted A4' and is used in all the motion resolution systems described henceforth in this thesis.

A4' The new observer absolute state estimate is added to the latest target relative state estimates to give new estimates of the target absolute states.

This modification is especially attractive for motion resolution because it makes the algorithm cheaper to compute.

The major deficiency in the experimental implementations is their lack of a proper state transition noise model. The targets and observer were assumed to move deterministically, and so the transition noise covariance matrix was identically zero for all targets and for the observer. Deviation from deterministic observer behaviour (for example in tests 3 and 4) caused random biassed acceleration components in the target relative state vectors. The deterministic motion model is invalid for marine targets in general, because of the perturbation caused by unknown small-scale water currents, and a better founded error model is required.

Apart from its theoretical deficiency, the deterministic motion model is numerically unsatisfactory since it allows the Kalman filter gains to converge asymptotically to zero and encourages numerical instability in the filters. The latter was one of the reasons why the measurement model noise variance had to be larger than the true measurement error variance (the second contributory factor was the numerical instability of the Kalman gain formula itself).

The asymptotically decreasing gain also causes the Kalman filters to place increasing confidence in their state estimates at the expense of making best use of new measurements. This effectively increases the response time of the recursive cycle and encourages the persistence of transients caused by targets changing their direction or acceleration.

The test implementations assumed that the number of targets was known at initialisation time and that the mapping between measurements (echoes) and the targets (echo sources) was known. In a practical system, the number of targets being tracked varies with time as new targets are observed or old ones move out of range, and the mapping from an event (which is anonymous) to its source is non-trivial to determine. Thus a practical motion resolution system, while still consisting of a realisation of steps A1, A2, A3 and A4' of algorithm A, requires a more flexible architecture than the preliminary implementations possessed.

There remains the possibility of using the spatial correlation of measurement noise in the estimation. The test implementations neglected this effect and were unable to take account of the fact that the range uncertainty in the test measurements (±3 centimetres) was substantially less that the angular uncertainty (±1 metre, roughly). This correlation is known from the range and azimuth values of the measurement (see Figure 4.3) and is a profitable source of information for the system.

The preliminary experiments therefore showed that algorithm A, suitably modified, was suitable for performing motion resolution. The initialisation scheme of 4.1.5 was able to function with two moving targets in test 2 and the system performed well in test 3 and 4 despite its inadequate noise model.

The outstanding issues are the definition of an adequate state transition noise model, the design of an architecture appropriate to the practical motion resolution problem, and the exploitation of the correlation between measurement noises. These problems are addressed by the new, practical motion resolution system design described in the next section.

4.3 A Practical Two Dimensional Motion Resolution System.

Using the insights obtained during the evaluation of the test implementations, a new, practical, linear motion resolution scheme based on steps A1, A2, A3 and A4' of algorithm A was designed. This scheme includes an explicit model of target motion from which state transition noise matrices may be calculated for the Kalman tracking filters. It has an architecture that reflects the computational structure of the motion resolution problem.

4.3.1 System Organisation.

Motion resolution processing divides naturally into two parts: those computations that maintain observer position and velocity estimates, and those that maintain the relative and absolute position and velocity estimates for each target. The former computations are not specific to any target, while the latter are duplicated once for each currently known target.

The target-specific computations comprise the tracking of input relative position measurements to provide apparent motion estimates (step A1), the computation of observer motion estimates based on the information available from a single target (step A2) and the maintenance of target absolute position and velocity estimates using target relative motion and observer motion data (step A4').

Observer-specific computations comprise the collation of observer motion estimates provided by individual targets, their composition into observer motion estimates incorporating information from many targets, and the maintenance of observer motion parameter estimates using these composite data. This computation is represented by step A3.

The practical motion resolution system reflects this division into target-specific and observer-specific functions: it is constructed as a kernel, comprising the observer-specific processing, supporting a number of target channels. Each channel provides the target-specific

processing for a single target.

4.3.2 System Operation.

The schematic structure of the motion resolving system is shown in Figure 4.5. The kernel and a single channel are shown; the boundary between the two is marked with a dashed line.

The system operates as follows. Each time a particular target is observed, the event vector is passed to the appropriate channel and the target relative state (TRS) vector, $\hat{\underline{x}}$, for that target is updated by the relative motion tracking filter, Kr. (The implicit problem of matching events to targets is handled by the segmentation module and is the subject of Chapter five.) If the target is a new one (not previously seen) the TRS vector of a new channel is initialised using the event vector, and the error covariance matrix is initialised from the measurement error covariance. The channel receives five observations, allowing the TRS estimate to settle, without contributing to the observer estimation process; each subsequent observation leaves the channel in an active state.

The activation of a channel indicates that it requires kernel servicing. The channel continues to update its TRS vector if events arrive during its activation, and it remains in the active state until deactivated by the kernel.

Whenever there are activated channels outstanding, the kernel processing estimates the observer motion parameters, requesting individual observer motion estimates from the active channels. Channels active for the first time are ignored during this step. Channels that have been activated at least twice contribute two observer motion estimates to the kernel:

1) an estimate of observer position calculated by subtracting the new target relative position from the target absolute position, computed at box 1 in Figure 4.5;

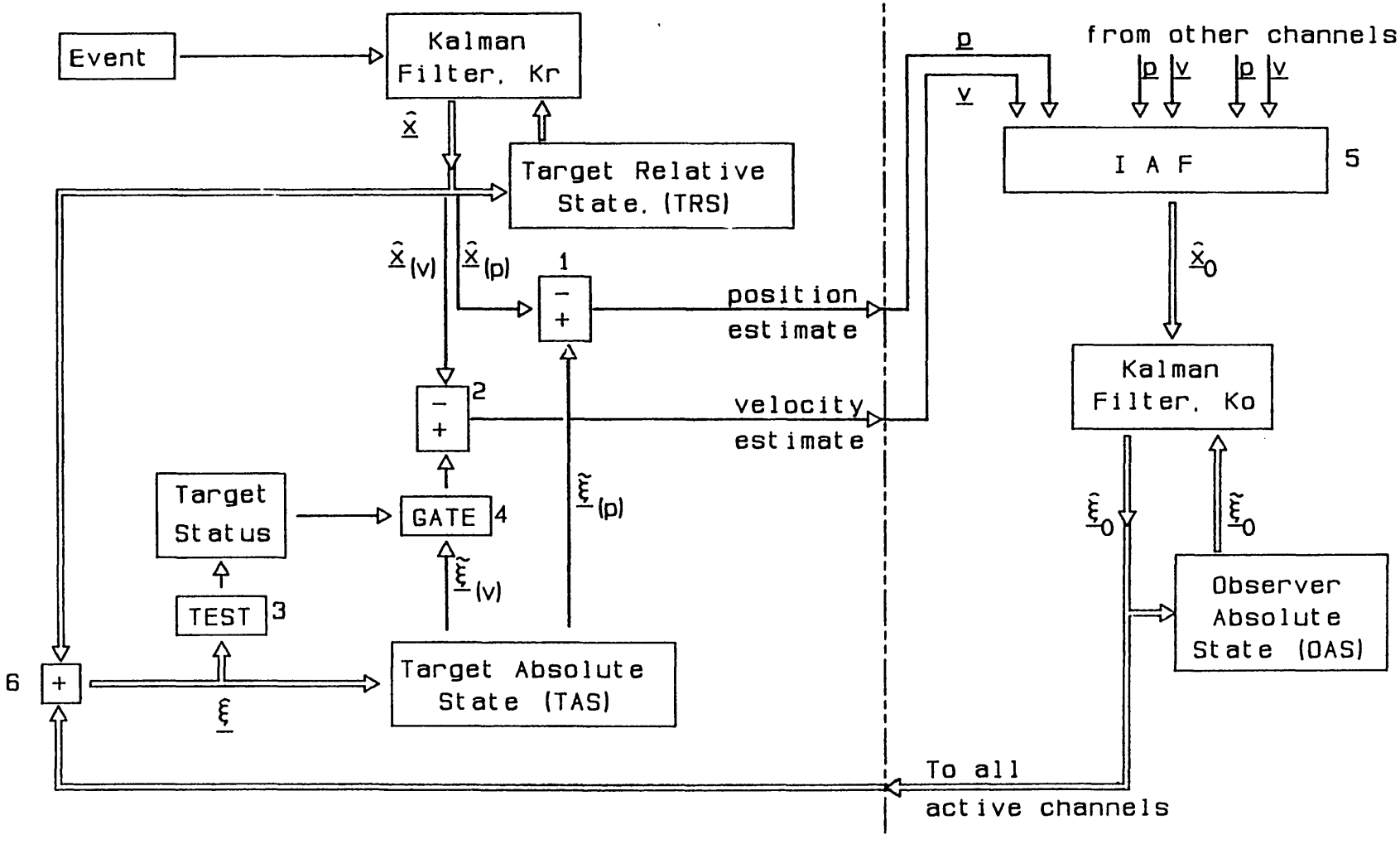


Figure 4.5. The New Motion Resolution System

2) an estimate of observer linear velocity, obtained from the target relative velocity corrected if necessary for target proper motion. This is computed by box 2 in the figure.

The target absolute velocity estimate is gated into this computation only if the target status resulting from the previous hypothesis test (box 3) indicates that the target is a moving one. The gate, box 4, effectively sets to zero the absolute velocity of targets currently hypothesised to be stationary.

The kernel collates these estimates and computes composite observer motion estimates (in $\hat{\underline{x}}_0$) using an Information Averaging Filter, shown as box 5. The composite estimates are tracked by the observer tracking filter, Ko, generating a new estimate of the observer absolute state (OAS) vector.

The new OAS vector, $\hat{\underline{\xi}}_0$, is passed back to all the currently activated channels, which use it to compute a new target absolute state (TAS) vector, $\hat{\underline{\xi}}$, (via box 6), containing target position and proper velocity estimates. For a channel activated for the first time this operation initialises the TAS vector. The target status (stationary or moving) is set according to the result of a stationary target hypothesis test applied to the new TAS vector by box 3. The channels which received the new OAS are then deactivated by the kernel.

4.3.3 The Target Motion Model.

The model of target motion, from which the state transition noise covariance matrices for the motion resolution scheme Kalman filters are computed, rests on the assumption that each target moves along a randomly perturbed piecewise linear trajectory. The random perturbation is modelled by an additive vector sequence obtained by integrating a Gaussian acceleration using an autoregressive process. The perturbation is isotropic and uncorrelated among the spatial dimensions.

The noise model describes the errors to be expected in the Kalman tracking filter predictions of target position and velocity. It implies that the difference between these state predictions and the actual target states is given by a random perturbation vector of the sort described above. This model, which assumes a random acceleration perturbation, matches the effects of the marine environment more closely than the deterministic model used in the previous experimental implementations (where it was assumed that the state predictions generated by the Kalman filters agreed precisely with the corresponding actual target states). Large scale water currents are modelled as piecewise linear velocity bias, while small scale currents are handled by the random perturbation component. The model does not deal with periodic water currents.

The position and velocity noise sequences are defined recursively in terms of the sample time quantum τ, a unit variance zero mean Gaussian acceleration sequence $\{a_n\}$, and the acceleration variance A^2:

$$\underline{s}_n = \phi(1)\underline{s}_{n-1} + (A\tau a_n)\underline{q} \qquad (\underline{s}_0 = 0) \tag{4.17}$$

where

$$\underline{s}_n = \begin{bmatrix} s_n \\ v_n \end{bmatrix}, \qquad \underline{q} = \begin{bmatrix} \frac{\tau}{2} \\ 1 \end{bmatrix} \quad \text{and} \quad \phi(n) = \begin{bmatrix} 1 & n\tau \\ 0 & 1 \end{bmatrix}. \tag{4.18}$$

In this equation, the term $(A\tau a_n)\underline{q}$ contains the position and velocity increment induced by the random acceleration Aa_n acting over the interval τ, and $\underline{s}_n$ is the accumulated position and velocity perturbation of all accelerations up to, but not including, a_n.

With the definitions (4.17) and (4.18) the covariance matrix of the sequence is

$$\begin{aligned} Q(n) &= \phi(1)Q(n-1)\phi(1)^T + (A\tau)^2 \underline{q}\,\underline{q}^T \qquad (Q(0) = 0) \\ &= n\tau^2 A^2 \begin{bmatrix} (4n^2-1)\frac{\tau^2}{12} & n\frac{\tau}{2} \\ n\frac{\tau}{2} & 1 \end{bmatrix}. \end{aligned} \tag{4.19}$$

The $Q(n)$ matrix is the covariance of the state transition noise in a single spatial dimension after an interval $n\tau$ has elapsed, just as $\phi(n)$ is the state transition matrix for a single dimension for that interval.

In the previous experimental implementations, which assumed that there was no random perturbation of target motion, the tracking filters estimated position velocity and acceleration for each target. The improved target motion model does not require that filters estimate target acceleration, since this is encompassed by the random perturbation vector in the model. Filter state vectors in the new system therefore include no acceleration components.

The state vectors actually used in the new system have four components (the estimates of position and velocity in each spatial dimension) which enable the filters to estimate the cross correlation between spatial dimensions. The state transition noise matrix Q_{mn} for these filters is a four by four matrix defined in terms of the components of $Q(k)$, where k is chosen so that $k\tau=t_n-t_m$. If the state vector is set out $[x_{pos}\ y_{pos}\ x_{vel}\ y_{vel}]^T$ and the components of $Q(k)$ are $\underline{q}_{pp}$, $\underline{q}_{pv}$ and $\underline{q}_{vv}$ for position variance, cross-covariance, and velocity variance, then the state transition noise matrix is

$$\begin{bmatrix} \underline{q}_{pp} & 0 & \underline{q}_{pv} & 0 \\ 0 & \underline{q}_{pp} & 0 & \underline{q}_{pv} \\ \underline{q}_{pv} & 0 & \underline{q}_{vv} & 0 \\ 0 & \underline{q}_{pv} & 0 & \underline{q}_{vv} \end{bmatrix}.$$

This matrix describes an isotropic perturbation (since corresponding entries for x and y are equal) and there is no cross-correlation between spatial dimensions (all the correlation terms in the matrix are zero).

The transition noise vector $_i\underline{q}_{mn}$ for each target is constructed by inserting the components of two $\underline{s}_k$ vectors appropriately into the four component vector.

4.3.4 Initialisation.

There are two sets of data requiring initialisation in the improved design: the observer motion parameters in the kernel, which are initialised once, and the target motion parameters, which must be initialised when a new channel is created. The kernel parameters are initialised using the position and velocity estimates computed from the first set of activated channels (a channel does not become active until its own initialisation has completed and its target has been sighted again).

The relative tracking filters (Kr) in new channels are initialised using a measurement event vector, and are allowed a further four measurements so that the relative tracking filter transients die away before the target relative state vector is included in the observer motion estimation. (Since the filters do not track acceleration they require less settling time than those in the experimental implementations.) Once the five measurement start-up phase is finished each subsequent observation activates the channel. Initial relative position estimates are obtained from the measurement responsible for the channel's creation, while the initial relative velocity estimate is just the current estimate of the reflected motion field for the new target (using the assumption that all targets are stationary until proven otherwise). Target absolute states are initialised using the procedures discussed in section 4.1.5 as soon as the channel is first activated.

The transition noise model provides a mechanism for initialising the error covariance matrix in a new channel. In the limit, as $\tau \to 0$, $n\tau \to t$, and $A^2\tau \to \alpha$, the matrix $Q(n) \to Q(t)$ where

$$Q(t) = \alpha t \begin{bmatrix} \frac{t^2}{3} & \frac{t}{2} \\ \frac{t}{2} & 1 \end{bmatrix}. \qquad (4.20)$$

This matrix gives the error covariance of the noise vector obtained from the sequence $\underline{s}_n$ at the end of a time interval t as a continuous function of the elapsed time.

To initialise a new channel error covariance, three values are required per spatial dimension: a position error variance, a velocity error variance, and a position-velocity covariance. The first of these is available from the measurement event vector. The other two are computed using the relationships between the components of the $Q(t)$ matrix. If the position covariance obtained from the measurement is M then the initial channel error covariance is given, as a partitioned matrix, by

$$P_{init} = \begin{bmatrix} M & \frac{3M}{2\tau_i} \\ \frac{3M}{2\tau_i} & \frac{3M}{\tau_i^2} \end{bmatrix}, \quad \text{where} \quad M = \begin{bmatrix} \sigma_{xx}^2 & \sigma_{xy}^2 \\ \sigma_{xy}^2 & \sigma_{yy}^2 \end{bmatrix}. \tag{4.21}$$

This is obtained by scaling the matrix $Q(\tau_i)$, where τ_i is a parameter with the dimensions of time, so that its position variance component is equal to $\sigma_{xx}^2, \sigma_{xy}^2$ and σ_{yy}^2 in turn, and inserting the components of the scaled $Q(\tau_i)$ appropriately in to the error covariance matrix. The parameter τ_i determines the uncertainty in velocity assumed by the initialisation. To see its significance, assume that the velocity distribution of targets seen for the first time is uniform and that $v \in [-V_{max}, V_{max}]$. The variance of v is then given by

$$\sigma_v^2 = \frac{V_{max}^2}{3} = \frac{3\sigma_p^2}{\tau_i^2},$$

where σ_p^2 is the variance of the position measurement, whence

$$V_{max} = \frac{3\sigma_p}{\tau_i}.$$

Thus the initialisation time τ_i is determined by the maximum velocity of first time targets.

4.4 Performance Tests on the New Design.

The performance of the new design was evaluated using simulation tests of a similar type to those described in sections 4.2.1 and 4.2.2. Four series of tests were run; they are described in detail in sections 4.4.2 to 4.4.5 respectively. For each test the error differences between the true absolute state (known to the simulator) and the estimated absolute state (maintained by the motion resolution algorithm) were accumulated for the observer and each target. These statistics form the basis of the performance assessment.

The algorithm's performance in any given test was evaluated with respect to five categories. These, and the statistics providing evidence in each category of assessment, are described next.

The first category is velocity extraction, or the ability of the algorithm correctly to estimate the absolute velocities of targets and observer. This aspect of performance can be assessed directly from the mean error velocity vector, and indirectly from the position error variances, for targets or observer. The position error variances, if small, indicate that any apparent velocity bias (suggested by the mean error vector) has not caused a substantial position error over the test period (the position of a target is not simply the integral of its velocity because of the Kalman filter action of the algorithm: each new measurement causes a position correction to occur).

The second category is noise reduction. This can be assessed directly by comparing the position error covariances for the absolute state components of a target with the corresponding covariance of the measurement noise for the target.

Third, the initialisation performance of the algorithm is assessed. This aspect of performance is known indirectly from the overall success or failure of the algorithm, and directly from the hypothesis test decisions -- these show whether the algorithm is in an initialisation phase or is operating normally. Some indirect evidence is also available from the observer mean position error since

(in all the tests described) this is due mostly to origin drift during the filter initialisation and settling period. This aspect of performance is of particular interest where there are moving targets since the algorithm will, in general, then start up in an initialisation phase.

The fourth category is temporal stability, or the stability of the viewpoint independent reference frame over time. This is clearly related to velocity extraction in that if the velocity extraction performance is good there is little residual velocity to cause temporal drift in the reference frame position. The temporal stability of the reference frame is indicated by the observer mean velocity error and position error variance (which describe the error motions of the reference frame origin), with the same comments about the variances as for velocity extraction.

Finally, the fifth category of assessment is spatial stability, or the accuracy of the computed reference frame. The errors in motion resolution cause the estimated target positions to be displaced with respect to their true positions. If the targets are imagined to be attached to an elastic sheet, then the effect of moving the targets from their true positions, where the sheet is undistorted, to the estimated positions is to introduce a distortion into the elastic. This is the idea underlying spatial stability -- it is an assessment of the metric distortion introduced by motion resolution. The distortion can be estimated from the position error offsets by inspection, as in section 4.2.2, or from the true and estimated positions using the affine transformation analysis technique described below.

4.4.1 Affine Transformation Analysis.

So far, the spatial distribution in the viewpoint independent reference frame constructed during a test has been assessed using a subjective judgement based on the absolute position offset values for the observer and targets. If the offset vectors are similar then the frame is judged to be spatially consistent. Dissimilar offsets are indicative of spatial distortion.

This subjective judgement is difficult to quantise and thus it is hard to compare performance in different tests using this method. This section presents an analytic method for assessing spatial distortion -- Affine Transformation Analysis -- which gives a direct numerical indication of both the nature and the degree of the distortion present in the viewpoint independent reference frame.

Affine transformation analysis measures distortion by computing an optimal (minimal square error) affine transformation between the set of true target absolute positions and the estimated absolute positions provided by motion resolution. The situation is illustrated in Figure 4.6, for the example given below. The true positions $\underline{x}_i$ are mapped by a rotation and stretch, then a shift of origin, into the estimated position $\underline{y}_i$.

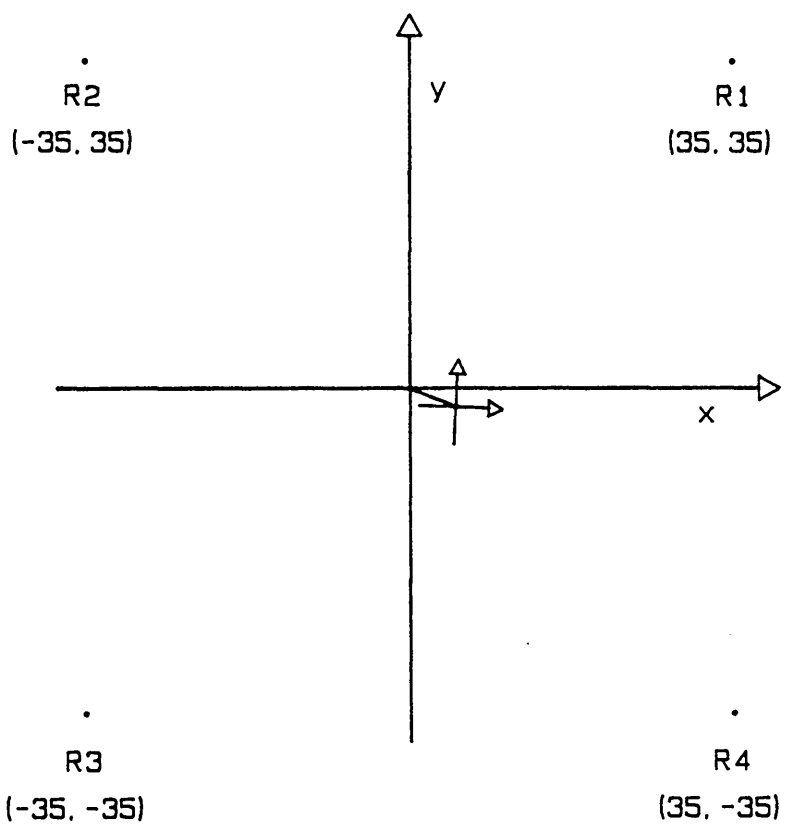


Figure 4.6. Mapping from True to Estimated Positions.

The general affine transformation is defined by two parameters, a vector $\underline{d}$ describing the origin shift and a matrix T describing the rotation and dilatation part of the transformation. The true and

estimated positions are related by

$$\underline{y}_i = T\underline{x}_i + \underline{d}, \quad \text{for } i=1..N, \tag{4.22}$$

and the optimal transformation is chosen by minimising the total error residual outer product

$$E = \sum_{i=1}^{i=N} (\underline{y}_i - T\underline{x}_i - \underline{d})(\underline{y}_i - T\underline{x}_i - \underline{d})^T \tag{4.23}$$

as a function of $\underline{d}$ and T. The minimal value of E, a matrix, indicates the quality of the fitted transformation.

Given the value of T, it is easy to show that

$$\underline{d} = \frac{1}{N} \sum_{i=1}^{i=N} (\underline{y}_i - T\underline{x}_i) = \underline{y} - T\underline{x} \tag{4.24}$$

where $\underline{x}$ is $\frac{1}{N} \sum_{i=1}^{i=N} \underline{x}_i$ and $\underline{y}$ is $\frac{1}{N} \sum_{i=1}^{i=N} \underline{y}_i$. With this value of $\underline{d}$, the matrix T is given by

$$T = \{ \sum_{i=1}^{i=N} (\underline{y}_i - \underline{y})(\underline{x}_i - \underline{x})^T \} \{ \sum_{i=1}^{i=N} (\underline{x}_i - \underline{x})(\underline{x}_i - \underline{x})^T \}^{-1}. \tag{4.25}$$

(This is easily proved by multiplying out (4.23) and substituting $T+\beta$, where T is given by (4.25), for the T in (4.23); β is an arbitrary matrix of suitable dimension.)

If the matrices X, Y and M are defined by

$$X = \{ \sum_{i=1}^{i=N} \underline{x}_i \, \underline{x}_i^T \} - N\underline{x} \, \underline{x}^T, \tag{4.26a}$$

$$M = \{ \sum_{i=1}^{i=N} \underline{y}_i \, \underline{x}_i^T \} - N\underline{y} \, \underline{x}^T, \text{ and} \tag{4.26b}$$

$$Y = \{ \sum_{i=1}^{i=N} \underline{y}_i \, \underline{y}_i^T \} - N\underline{y} \, \underline{y}^T, \tag{4.26c}$$

then the transformation matrix T and error matrix E are given by

$$T = MX^{-1} \quad \text{and} \quad E = Y - MX^{-1}M^T. \tag{4.27}$$

The data for an example of affine transformation analysis is given in Table 4.7. The actual transformation between the true and estimated positions was a rigid rotation of 50 milliradians and an origin shift of (0.5,-0.2), but the transformed positions were also perturbed using small vectors (components in the range ±0.2 metres).

	True Position/m		Estimated Position/m	
Target	x	y	x	y
1	+35	+35	37.305	32.857
2	-35	+35	-32.757	36.496
3	-35	-35	-36.026	-33.207
4	+35	-35	33.507	-36.866

T		d/m	E/m^2	
0.9971	0.0505	0.5073	0.06996	0.00252
-0.0521	0.9959	-0.18	0.00252	0.0001

Table 4.7. An Affine Transformation Analysis Example.

The results of the analysis are also given in Table 4.7. The matrix T is close to the rigid rotation matrix used to transform the true data, and the displacement vector $\underline{d}$ agrees reasonably with the actual origin shift used. The small values in the E matrix indicate the high quality of the fit of the affine transformation.

In the results tables given later in this chapter and in Chapter five the affine analysis data is further condensed. The displacement vector is omitted, since it represents the initialisation phase drift of the (initially arbitrary) viewpoint independent reference frame and so is of little interest. Two indices of the transformation matrix are given: the average rotation and the expansion factor. The former is the angle whose tangent is the ratio of the difference of off-diagonal elements to the trace of the matrix. If the transformation is a rigid rotation (and the results of the affine analysis are generally very close to a rigid rotation) then this angle is just the rotation angle of the matrix. The expansion factor ε is equal to **det**(T) - 1. If T is a rigid rotation, then **det**(T) is equal to unity and the expansion factor is zero. For transformations that include a dilatation, the determinant of T is not unity, and the expansion factor ε indicates the incremental change in area induced by the transformation. The eigenvalues of the matrix $\frac{1}{N}E$, where N is the

number of targets included in the affine analysis, are also given. These values are the maximum and minimum variances of the average error of the transformation fit.

For perfect performance of the motion resolving system on perfect data, the estimated positions should be related to the true ones by an arbitrary displacement only (in Chapter five, an angular displacement is also to be expected). Thus the expansion factor should be zero, the eigenvalues of E small, and the average rotation should also be zero.

In fact, the data is not perfect because of measurement noise and so the affine analysis indices are slightly corrupted. In the example above, the average rotation is 51.4 milliradians approximately, which is 2.8% higher than the true value; the expansion factor is 3.067%, indicating that the transformation induces a small expansion; and the fit error eigenvalues are 0.07 m^2 (the maximum) and 1.64 mm^2 (the minimum), showing that the fit error is strongly direction dependent.

4.4.2 Linear Observer Motion Tests.

The first set of tests run on the new motion resolving system comprised three tests in which the observer moved along a linear trajectory with velocity (0.08,0.06) amongst four stationary targets. The four targets were arranged at the corners of a square centred on the origin, as in the tests described in section 4.2.2. In these tests, named the LIN series, the size of the square (the world size) was varied between the tests, taking the values 40 metres, 70 metres, and 100 metres respectively in lin1, lin2 and lin3.

Each test ran for 150 seconds of simulated time. The sonar repetition rate was 2.5Hz and the measurement noise parameters were set at ±3 cm range resolution and 3.4° azimuth resolution as before.

The observer and target motion statistics for the three tests in this series are presented in Table 4.7. Note that in all the tests the true observer motion was (80,60) in the units used in the table.

The mean velocity error ($\delta\underline{v}$) and the position error variance (PEV) are tabulated for each target in each of the three tests.

		Velocity Error $\delta\underline{v}/mms^{-1}$		PEV/cm^2	
Test	Target	x	y	x	y
	Obs	0.279	0.380	3.82	8.12
	1	-0.191	0.490	69.49	102.9
lin1	2	1.501	-1.147	94.44	96.76
	3	0.550	1.603	72.90	62.67
	4	-1.724	1.453	75.77	69.84
	Obs	0.111	0.254	2.21	9.28
	1	1.474	1.696	193.8	215.7
lin2	2	1.250	-0.511	217.7	222.0
	3	0.951	1.585	145.6	129.2
	4	-1.354	1.284	163.6	154.6
	Obs	0.177	0.047	3.30	7.41
	1	1.991	1.860	345.0	378.9
lin3	2	1.752	-1.108	371.4	370.3
	3	2.111	2.477	226.6	209.5
	4	-1.929	1.668	307.2	279.6

Table 4.7. Observer Motion Statistics for the LIN tests.

In terms of velocity extraction the system performed well. The observer velocity was estimated to better than 0.4 mms^{-1} in all the tests. This is an error of about 0.6%. The estimates of target velocity show slightly more error, but that is expected since the observer estimates incorporate data from all the targets (and so have a better signal to noise ratio than any individual target's data).

The position error variances, which indicate the temporal stability and noise reduction performance of the system, are also small. The worst case observer position error has a standard deviation of just over 3 cm, and the typical position error for a target has a standard deviation of under 9 cm for lin1, about 14 cm for lin2, and about 18 cm for lin3. These compare favourably with the measurement error standard deviations (mostly due to azimuth measurement noise) of roughly 28 cm for lin1, 50 cm for lin2, and 1 metre for lin3, giving variance reduction factors from just over 4 to about 31.

The spatial consistency of the reference frames generated in these tests is shown by the affine analysis indices, presented in Table

Test	Rotation/mrd	ε/%	λ_{min}/mm^2	λ_{max}/mm^2
lin1	0.127	0.01	8.656	108.0
lin2	0.179	0.01	42.74	351.7
lin3	0.174	0.01	99.75	735.3

Table 4.8. Affine Analysis Indices for the LIN tests.

4.8. In all the tests the affine transformation is very close to a rigid rotation (ε is very small) and the angle of that rotation is also small. Thus the reference frame generated by the motion resolving system corresponds closely to the world reference frame defined by the simulator program. The fit error variances are also small, corresponding to standard deviations of between about 3 mm and 27 mm. Note that the maximum eigenvalue is significantly larger than the minimum: this is due to the anisotropy of the measurement noise, which comprises small radial errors but substantial errors in the azimuthal direction. The stability of the reference frame, as indicated by these results, is very good indeed.

Conclusions for the LIN Tests.

The new motion resolution system is able to construct and stabilise a viewpoint independent reference frame in these simple tests. The observer velocity is extracted with a maximum error of about 0.6%, the observer position is estimated with an error of less than 3 cm standard deviation while target positions are estimated with a worst case error of 18 cm standard deviation. The temporal and spatial stability of the reference frame is good. No initialisation problems were experienced in these tests (but the system did not start up in an initialisation phase since there were no moving targets to classify).

4.4.3 The SL Series of Monte Carlo Tests.

The second set of tests run using the motion resolution system of section three comprised 15 Monte Carlo tests involving a moving observer and several stationary targets. The number of targets, the target positions, and the observer motion were all chosen at random for these tests. The series was named SL, denoting stationary targets and linear observer motion.

The number of targets was chosen uniformly between two and ten; the position of each target was specified by two coordinates selected at random from uniform distributions over the interval $[-100,+100]$. The observer moved along an unperturbed linear trajectory, and the velocity components for the observer were zero mean Gaussian random numbers with a standard deviation of 10 cms^{-1}. Each test ran for 120 seconds of simulated time, and the sonar noise model was the same as for the LIN tests, i.e. ±3 cm range resolution and about 3.4° beamwidth, with a 2.5Hz sampling rate.

		True Velocity/mms^{-1}		Mean $\delta\underline{v}/mms^{-1}$		PEV/cm^2	
Test	N	x	y	x	y	x	y
sl1	3	-30.54	31.43	1.510	0.561	11.28	0.84
sl2	5	88.20	85.30	0.962	-0.636	7.98	2.63
sl3	8	-5.64	32.99	0.419	-0.252	1.77	1.51
sl4	8	-56.93	76.51	0.319	0.149	3.45	1.39
sl5	3	-93.03	-4.08	-0.197	0.523	0.68	6.35
sl6	4	8.48	1.33	0.233	0.509	0.87	1.66
sl7	5	-19.54	-11.74	0.261	0.017	4.28	1.57
sl8	4	94.04	-9.75	-0.439	-0.159	1.60	3.83
sl9	5	-45.72	-3.49	0.482	0.307	2.68	0.53
sl10	4	-268.8	-149.4	1.383	1.493	16.82	53.43
sl11	4	144.7	152.9	1.270	-0.679	26.89	7.58
sl12	9	42.69	101.6	0.751	-0.359	8.52	1.22
sl13	10	-14.23	2.17	-0.167	0.100	1.38	0.71
sl14	6	-138.9	-4.50	-0.535	0.533	1.28	2.88
sl15	9	-100.0	38.79	0.668	0.398	3.37	3.22

Table 4.9. Observer Motion Data for the SL Test Set.

Table 4.9 gives the observer motion statistics for the tests in this series. For each test the true observer velocity, the mean velocity error, and the observer position estimation error variance

Test	$\delta v/mms^{-1}$ Max. & Min.	Approx. Range/m	Measurement Variance/m^2	PEV/m^2 x	y
sl1	7.50 (max)	100	1.00	0.08415	0.03305
	0.30 (min)	84	0.71	0.04958	0.03055
sl2	5.74	93	0.86	0.05841	0.01831
	0.08	44	0.19	0.03532	0.00051
sl3	9.56	108	1.17	0.09580	0.07480
	0.07	68	0.46	0.04520	0.00225
sl4	13.42	109	1.19	0.09039	0.07857
	0.36	68	0.46	0.05086	0.00093
sl5	7.92	70	0.49	0.00159	0.05747
	0.63	70	0.49	0.00159	0.05747
sl6	3.32	89	0.79	0.03181	0.05862
	0.25	96	0.92	0.08832	0.02003
sl7	5.43	82	0.67	0.03157	0.02415
	0.38	45	0.20	0.02221	0.03314
sl8	5.27	110	1.21	0.09299	0.06742
	0.14	29	0.08	0.00056	0.02318
sl9	9.69	107	1.14	0.07166	0.01894
	0.10	60	0.36	0.06952	0.00111
sl10	5.52	82	0.67	0.04095	0.02261
	0.38	116	1.35	0.1348	0.04271
sl11	7.92	100	1.00	0.1296	0.03340
	0.26	101	1.02	0.1089	0.01600
sl12	9.27	93	0.86	0.00361	0.1606
	0.06	124	1.54	0.07680	0.1030
sl13	14.86	93	0.86	0.00456	0.08910
	0.27	70	0.49	0.01305	0.05215
sl14	10.92	112	1.25	0.1719	0.03146
	0.64	43	0.18	0.02008	0.00174
sl15	8.39	101	1.02	0.00734	0.1164
	0.01	72	0.51	0.04956	0.01428

Table 4.10. Target Motion Statistics for SL Tests.

are tabulated. The velocity errors for these tests vary between 1.5 mms^{-1} and 0.02 mms^{-1}, with typical values being around 0.5 mms^{-1}. The typical relative error in these tests is about 1%, although some of the tests (with very small observer velocities, such as tests sl13, sl6, and to some extent sl1) show percentage errors as high as 38%. The position error variance is typically about 2 cm^2 and is almost everywhere less than 9 cm^2, corresponding to noise standard deviations of less than 3 cm and typically about 1.4 cm. The worst

case observer position error standard deviation is about 7.5 cm. Thus the reference frame is temporally stable and the velocity extraction performance is good.

The velocity extraction and noise reduction statistics for the targets in the SL tests are presented in condensed form in Table 4.10. This table lists for each test the maximum and minimum velocity error component (over all the targets in the test), and for the appropriate targets the approximate target range and measurement noise variance are tabulated together with the position error variance for the target. Note that the position error variance units are now m^2, ten thousand times the size of those in Table 4.9.

Test	Rotation/mrd	ε/%	λ_{min}/cm^2	λ_{max}/cm^2
sl1	0.036	0.03	0	6.703
sl2	-0.488	-0.04	0.8745	15.81
sl3	0.118	0.02	0.7253	4.496
sl4	-0.113	-0.01	3.082	6.321
sl5	0.544	0.03	0	4.117
sl6	-0.058	0.04	0.05486	0.6175
sl7	0.269	0.00	0.08944	12.23
sl8	0.161	-0.01	0.3009	2.286
sl9	-0.147	-0.04	0.983	2.989
sl10	-0.737	0.10	5.904	12.13
sl11	0.295	0.00	0.6174	16.48
sl12	-0.100	0.00	3.575	6.967
sl13	-0.295	0.02	2.488	4.195
sl14	0.057	0.00	1.187	5.834
sl15	-0.153	0.01	2.68	7.298

Table 4.11. Affine Analysis Indices for the SL Tests.

As before the system shows greater variation and less accuracy in its estimation of the state of the targets than in its observer motion estimation. The worst case mean velocity errors can be large, of the order of 1.5 cms^{-1}, with correspondingly large position error variances, but in general the target positions are estimated with an error standard deviation of about 30 cm and their velocities are typically estimated with between 1 and 5 mms^{-1} worst case error.

It is apparent from the table that the position error variances are correlated with the target range and measurement variance (the latter is proportional to the square of the range since the noise is

almost entirely due to the azimuth errors). In all the tests the system exhibits a significant noise reduction, the variance ratio of measurement to estimation noise being between about 4 and 30.

In Table 4.11 the affine stability analysis indices for the SL series tests are presented. The expansion factor is small for all the tests, showing good rigidity of the reference frame, and the fit errors are also small, typically 1 to 2 cm standard deviation, which indicates that the reference frame matches closely with the simulation world coordinate frame. There is a small angular displacement of the reference frame in most of these tests, generally of the order of 0.25 mrd, but in view of the relatively small numbers of points used in the affine analysis computation it is probably not significant (allowing for random variation it appears from the table that the average rotation tends to be greater for those tests with fewer targets).

Conclusions for the SL Tests.

The Monte Carlo tests give a sample of the system performance in the situation where the observer is moving among a number of stationary targets. There were no initialisation problems, not surprisingly, since there were no moving targets. The system was able to extract the observer linear velocity successfully with a typical accuracy of 0.5 mms^{-1} and estimated the observer position with an typical error of 7.5 cm standard deviation. The target velocity and position estimation errors were about four times the size of those for the observer, but the system showed significant noise reduction in all the tests. The temporal and spatial stability of the reference frame generated in the tests was good.

4.4.4 The ML Series of Monte Carlo Tests.

The third set of tests run on the linear motion resolution system comprised fifteen Monte Carlo tests in which both observer and targets were permitted to move. Each test included four stationary

targets arranged in a 70 metre square centred on the world origin, and between one and six moving targets. The series was named ML, denoting moving targets and linear observer motion.

The number of moving targets was chosen uniformly at random between one and six, and the target initial position coordinates were drawn from uniform distributions over the interval [-100,+100]. The observer moved along an unperturbed linear trajectory with velocity components given by zero mean Gaussian random numbers with 10 cms^{-1} standard deviation. Moving target velocity components were also zero mean Gaussian but had a 20 cms^{-1} standard deviation.

Each test ran for 120 seconds of simulated time and used the same sonar model as the LIN tests, i.e. an angular noise standard deviation of 10 mrd, a range resolution of ±3 cm, and a repetition rate of 2.5 Hz.

Table 4.12 gives the observer motion statistics for the fifteen tests in the series. For each test the table shows the true observer velocity $\underline{v}$, the mean velocity error $\delta\underline{v}$, and the observer position error variance (PEV). The number of targets in the test is tabulated as moving/total.

Of the fifteen tests, seven were deemed to have failed on the basis of their observer velocity extraction performance. They are listed in the first section of Table 4.12 and marked with an asterisk. In each of these tests the system misclassified all the stationary targets and therefore was unable to stabilise the absolute reference frame.

The cause of failure is illustrated in Figure 4.7, which shows as scatter plots the actual and estimated observer and target positions for test ml1. The four stationary reference targets, five to eight, drift throughout the test, and moving targets (for example two and four) also suffer velocity estimation errors as a consequence.

In eight of the tests the system correctly identified the stationary and moving targets. These successful tests show very good velocity extraction performance, with typical mean observer velocity

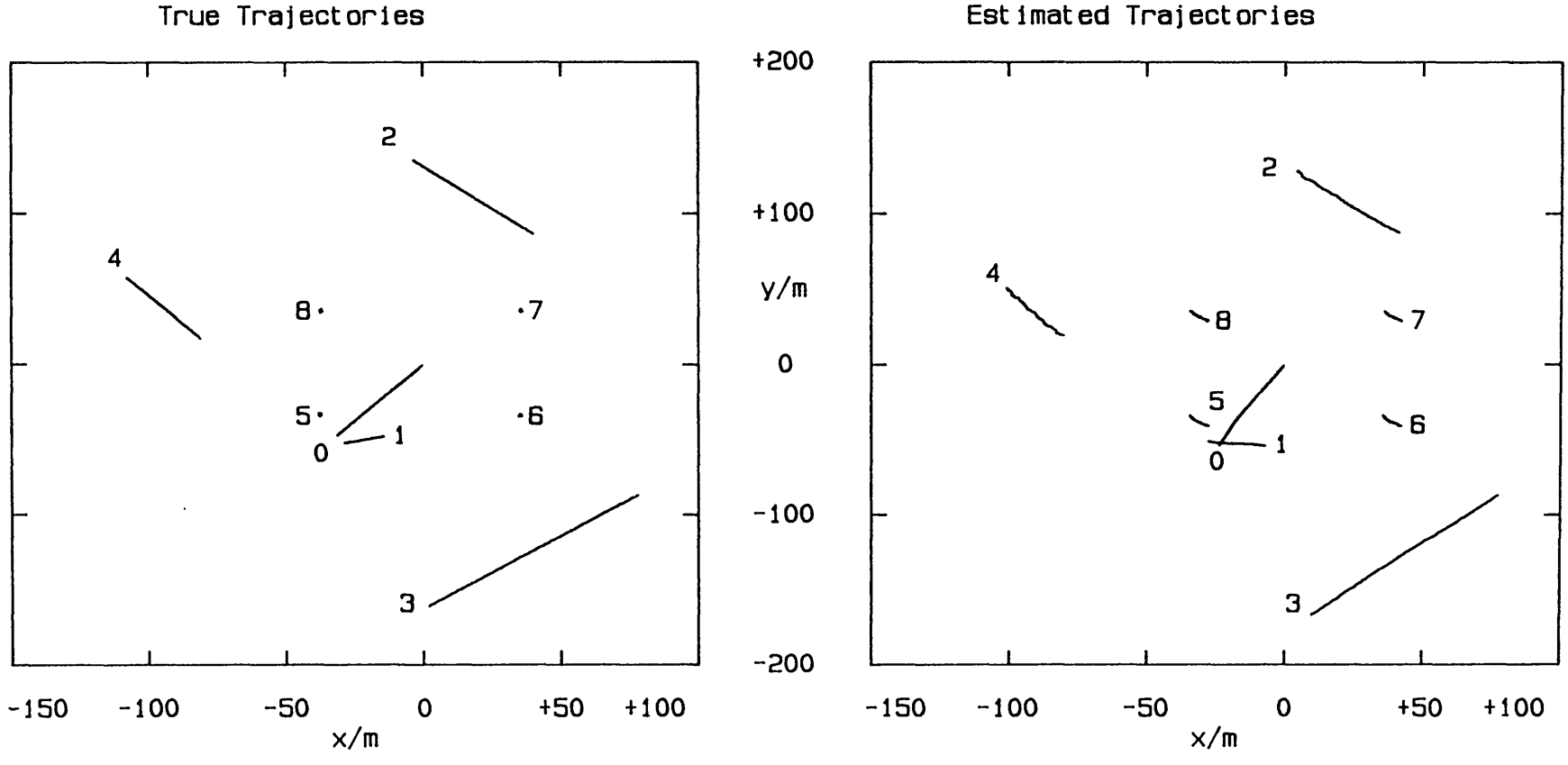


Figure 4.7. Scatter Plot of Absolute Target Positions for Test ml1.

errors of less than 0.5 mms^{-1}. The relative errors vary from 2.8% in test ml4 (where the y component of the true velocity is small) to 0.02% in test ml14, with typical values between 0.4% and 0.05%.

The position error variances for the successful tests show good noise reduction and temporal stability. Note that the units in the second section of Table 4.12 are 10,000 times smaller than those in the first section. The observer position is estimated with a worst case error of about 6 cm standard deviation; the typical position error in the successful tests had less than 3 cm standard deviation.

		$\underline{v}/mms^{-1}$		Mean $\delta\underline{v}/mms^{-1}$		PEV/m^2	
Test	N	x	y	x	y	x	y
ml1*	4/8	-262.8	-395.0	62.5	-58.9	3.176	2.503
ml2*	4/8	-387.9	222.9	48.0	230.8	2.450	40.59
ml6*	3/7	-131.2	106.5	-244.3	-46.3	45.37	2.632
ml7*	5/9	219.9	-362.4	-157.2	-68.6	21.53	3.774
ml10*	3/7	-126.0	-342.8	-255.3	187.2	56.68	30.93
ml11*	5/9	332.9	-125.1	17.49	-8.40	0.305	0.058
ml12*	5/9	31.54	51.83	-108.2	103.1	8.646	8.845
						PEV/cm^2	
ml3	3/7	223.7	594.8	-0.316	0.306	1.50	2.10
ml4	2/6	222.0	29.00	0.711	-0.808	6.60	11.50
ml5	4/8	-13.54	-82.21	0.215	0.335	10.86	12.32
ml8	5/9	-642.9	75.80	-0.477	0.567	1.66	3.09
ml9	2/6	-779.7	156.7	-0.573	-0.336	2.01	1.93
ml13	6/10	-389.5	273.3	0.265	0.054	7.70	2.64
ml14	5/9	-360.8	318.5	0.374	0.055	1.00	0.81
ml15	3/7	-117.2	-696.8	-1.71	-2.49	11.90	37.10

Table 4.12. Linear Observer Motion Data for the ML Test Set.

Table 4.13 presents condensed velocity extraction and noise reduction statistics for each of the ML tests deemed to have succeeded. The table comprises two parts. In each test the maximum and minimum mean velocity error component for moving targets were determined, and the target speed, approximate range, and position error variances were tabulated in the left hand section of the table. The maximum and minimum velocity errors and the corresponding position error variances were also determined for the four stationary targets in the test, and these values are listed on the right in the table. All the stationary targets were initially 50 m from the observer.

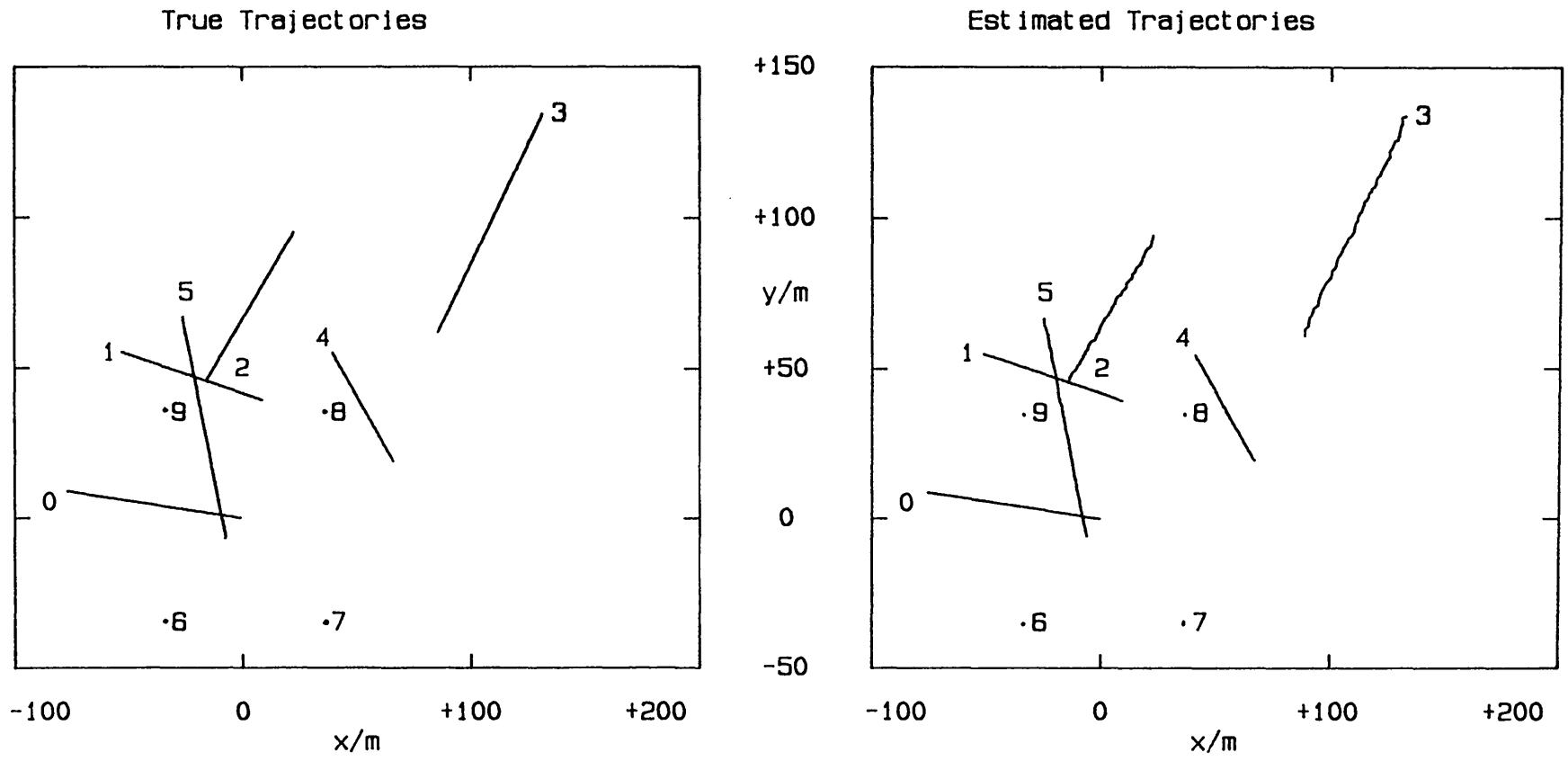
Test	N	Moving $\delta\underline{v}/mms^{-1}$	Moving $\lvert\underline{v}\rvert/mms^{-1}$	Moving Range/m	Moving PEV/m^2	Stationary $\delta\underline{v}/mms^{-1}$	Stationary PEV/m^2
ml3	3/7	17.18 (max)	1090	86	0.069	10.6	0.048
		3.22 (min)	567.7	67	0.012	0.11	0.003
ml4	2/6	3.48	130.7	39	0.027	14.1	0.012
		0.10			0.001	0.04	0.006
ml5	4/8	48.4	312.8	83	0.053	1.90	0.023
		0.20	168.4	97	0.003	0.75	0.012
ml8	5/9	151.4	723.7	118	0.141	6.49	0.042
		0.63	540.7	42	0.006	0.08	0.013
ml9	2/6	14.4	349.1	84	0.044	1.12	0.007
		3.37	195.4	59	0.024	0.05	0.004
ml13	6/10	16.3	384.2	82	0.097	5.91	0.043
		2.26	165.1	93	0.025	0.00	0.002
ml14	5/9	13.2	166.1	104	0.120	0.92	0.025
		1.26	518.4	121	0.006	0.01	0.003
ml15	3/7	63.4	356.2	94	0.180	65.6	0.019
		0.15	665.5	36	0.020	0.46	0.008

Table 4.13. Target Motion Statistics for the ML Test Set.

The moving target velocity estimates are less accurate than the observer velocity estimates since the latter incorporate information from all the targets whereas the former are based principally on the data for the target concerned. The worst case relative error is about 20%, in test ml8, while the best is 0.02% in test ml15. The typical relative error is better than 4% and the velocities of moving targets are estimated to better than 2 cms^{-1} in general, while stationary target velocity errors are typically 1 to 6 mms^{-1}.

The unusually high mean velocity error for a stationary target in test ml15, coupled with a small position error covariance, arose because that target was classified as moving by the system. The reduced performance apparent in this test is an indication of the performance of the motion resolving algorithm with three stationary targets instead of four.

The position error variances show reasonable temporal stability, the worst case error having between 42 and 16 cm standard deviation for moving targets, and the best having between 3 and 16 cm standard deviation. All the tests show target position noise reduction, with



Targets move towards their numbers; the observer is target 0.

Figure 4.8. Scatter Plot of Absolute Target Positions for Test m18.

the ratio of measurement to estimation error variance lying between 1.2 and 313 (the measurement noise variance is roughly 10^{-4} times the square of the range, since measurement noise is principally angular with 10 mrd standard deviation), and the typical noise reduction factor for the moving targets is about ten.

Stationary target positions are estimated with a worst case error of typically 15 cm standard deviation. The best case position error standard deviation varies between 4.5 cm (test ml13) and 11.4 cm (test ml15). The stationary targets in all the tests show noise reduction with worst case variance ratios between 10 and 42, and best case values as high as 250.

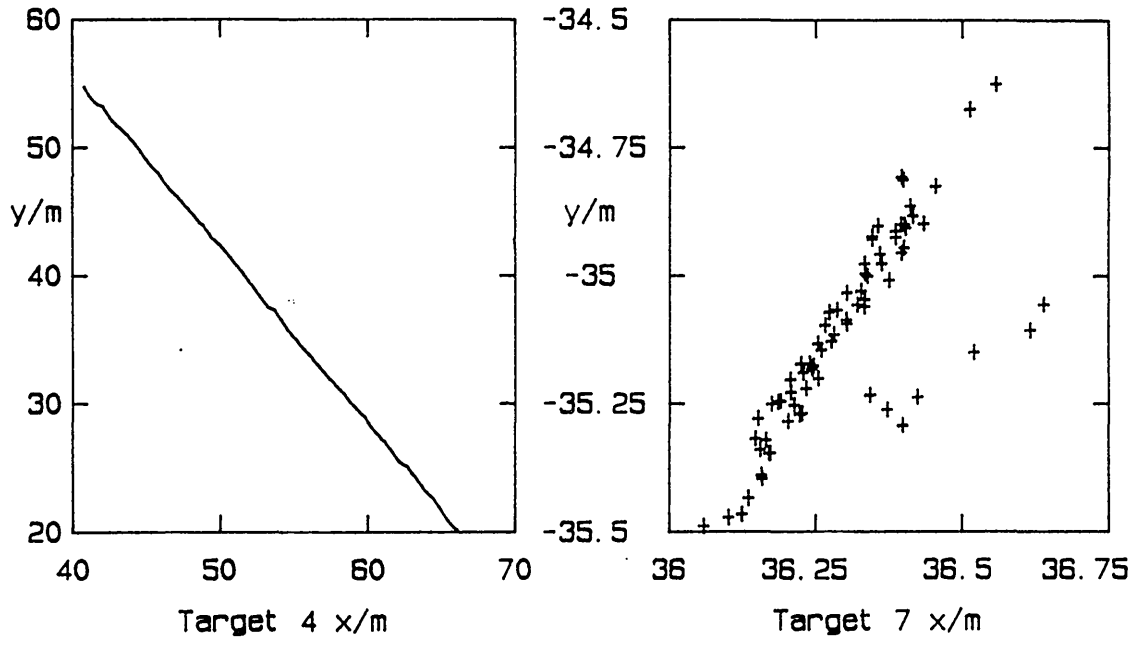


Figure 4.9. Tracks for Targets Four and Seven in Test ml8.

The tabulated data are complemented by Figures 4.8 and 4.9 which show the observer and target true and estimated absolute positions for test ml8 as x-y scatter plots. The scatter plots in Figure 4.9 are magnifications of the tracks of targets four (moving) and seven (stationary).

The stability of the viewpoint independent reference frame in a successful test is clear from the graphical presentation in Figure

Test	Rotation/mrd	ε/%	λ_{min}/cm^2	λ_{max}/cm^2
ml3	0.633	0.18	1.188e-03	1.646e-03
ml4	-0.149	0.00	0.74	2.52
ml5	-0.184	-0.03	1.69	27.13
ml8	-1.677	0.12	10.93	119.0
ml9	0.302	0.05	2.179	11.03
ml13	0.124	-0.10	17.95	1076
ml14	0.028	0.02	1.229	9.517
ml15	1.483	0.10	24.40	56.31

Table 4.14. Affine Analysis Indices for the ML Tests.

4.8, and this is confirmed by the affine analysis indices for the successful tests, which are presented in Table 4.14.

The unusually large affine fit error in test ml13 is also traceable to the testing of the target motion hypotheses. Of the four stationary targets in the test, one was correctly classified, but the remaining three were judged to be moving in more than half of the cycles of the motion resolution algorithm. All the moving targets were correctly classified. Thus in this test the viewpoint independent reference frame was established on the basis of a single stationary target with intermittent contributions from the other three.

Conclusions for the ML Tests.

The three major conclusions supported by the ML tests are these.

1. In the cases where the system initialisation is successful and the stationary targets are correctly identified (about 53% of the tests in this series) the viewpoint independent reference frame established by the system is spatially and temporally stable. The observer velocity extraction performance is good; target velocity extraction, although less accurate, is also satisfactory. The observer absolute position is tracked with better than 3 cm standard deviation, and the target absolute positions with typically 15 cm standard deviation. All the tests show noise reduction, with a variance ratio greater than ten.

2. In about 47% of the tests in the ML series the system failed to classify correctly any of the four stationary targets present. Since all the targets in these failed tests were judged to be moving, the velocity bias induced by the start-up initialisation phase persisted throughout the test. Consequently, the absolute reference frame origin drifted with time. The underlying cause of these failures was the poor performance of the target motion hypothesis test.

3. Test ml13 illustrates the performance achieved by the motion resolution algorithm when there is only one correctly classified stationary target on which to base the viewpoint independent reference frame. In view of the information available to the system in that situation, the performance shown in this test was adequate.

4.4.5 Perturbed Observer Motion: the NSE Tests.

The final series of tests on the linear motion resolution system comprised five tests based on test lin2 of sections 4.4.2. The arrangement of targets and observer, the sonar model parameters, and the deterministic observer motion were identical with test lin2. However, the five tests in this series had varying degrees of random perturbation superimposed on the linear observer trajectory. The degree of perturbation present increases uniformly from test nse1 through to test nse5. For this series the motion hypothesis testing mechanism was turned off in view of its poor performance in the ML series described above.

Table 4.15 shows the mean velocity error ($\delta \underline{v}$) and position error variance (PEV) for the observer, and gives the observer perturbation error mean (OPM) and variance (OPV) for each target in each test. The observer perturbation error is the difference between the actual observer motion and the deterministic trajectory defined by the observer velocity, (0.08,0.06).

Test	OPM/m^2 x	OPM/m^2 y	$\delta\underline{v}$/mms^{-1} x	$\delta\underline{v}$/mms^{-1} y	OPV/cm^2 x	OPV/cm^2 y	PEV/cm^2 x	PEV/cm^2 y
nse1	0.122	0.070	1.68	0.68	113.8	88.6	3.60	3.95
nse2	0.244	0.139	3.44	1.71	453.5	353.5	9.93	8.73
nse3	0.365	0.209	5.21	2.74	1020	795.1	20.90	16.89
nse4	0.487	0.279	6.96	3.77	1812	1413	36.53	28.40
nse5	0.609	0.348	8.73	4.80	2831	2208	56.81	43.27

Table 4.15. Observer Motion Statistics for the NSE Tests.

The observer motion statistics tabulated in Table 4.15 show good velocity extraction and temporal stability. The velocity extraction error is directly proportional to the mean value of the perturbation, suggesting that the system is accurately estimating the observer velocity. The observer position noise is reduced by a factor of 22 to 50 in all the tests. For tests nse2 to nse4, where the dominant noise source is the observer trajectory perturbation, the noise reduction factor is between 40 and 50.

Test	OPM/m x	OPM/m y	Target	Mean $\delta\underline{v}$/mms^{-1} x	Mean $\delta\underline{v}$/mms^{-1} y
nse1	0.122	0.070	1	1.441	1.306
			2	0.757	-0.678
			3	-0.096	0.319
			4	-0.939	0.205
nse2	0.244	0.139	1	1.171	1.033
			2	0.978	-0.743
			3	-0.475	0.104
			4	-0.751	0.156
nse3	0.365	0.209	1	0.912	0.756
			2	1.206	-0.812
			3	-0.870	-0.110
			4	-0.572	0.110
nse4	0.487	0.279	1	0.665	0.476
			2	1.443	-0.878
			3	-1.280	-0.320
			4	-0.406	0.061
nse5	0.609	0.348	1	0.433	0.193
			2	1.686	-0.940
			3	-1.697	-0.525
			4	-0.253	0.099

Table 4.16. Target Velocity Errors in the NSE Tests.

Test	OPV/cm² x	OPV/cm² y	Target	PEV/cm² x	PEV/cm² y
nse1	113.8	88.6	1	205.6	226.7
			2	215.4	221.1
			3	146.2	124.6
			4	180.0	168.2
nse2	453.5	353.5	1	222.3	237.2
			2	228.1	239.3
			3	149.8	130.1
			4	182.8	171.3
nse3	1020	795.1	1	242.7	250.8
			2	245.0	261.4
			3	156.5	139.2
			4	191.0	177.6
nse4	1812	1413	1	266.8	267.7
			2	266.0	287.2
			3	116.3	151.9
			4	202.4	187.1
nse5	2831	2208	1	294.8	288.0
			2	291.3	316.9
			3	179.1	168.2
			4	217.1	199.6

Table 4.17. Target Position Error in the NSE Tests.

Tables 4.16 and 4.17 give respectively the mean velocity errors ($\delta \underline{v}$) and the position error variances (PEV) for the four targets in each test. The mean perturbation (OPM) and the perturbation variance (OPV) are included in the tables for comparison. There are three points to make concerning these data.

First, it is evident from the tables that the target position and velocity estimates are affected less by the observer perturbation than the corresponding observer estimates. This indicates that the system is succeeding in resolving the apparent target motions.

Second, the position error variances increase gradually as the perturbation increases. The system noise reduction performance is good on the evidence of Table 4.17, with target position estimation errors having between 18 and 13 cm standard deviation in the presence of observer perturbations of 0.53 m standard deviation.

Third, the mean velocity errors also change as the perturbation increases. The estimation errors for targets 2 and 3 appear to rise roughly in step with the mean perturbation, but the errors for targets 1 and 4 fall in a similar manner. This latter trend is almost certainly connected with the fact that the observer is moving towards targets 1 and 4, and away from targets 2 and 3, but as yet I have no explanation for this.

To complement the tabular presentation of results in this section, Figure 4.10 shows the actual path of the observer during part of test nse5. The considerable random perturbation of the linear deterministic trajectory (which is shown as a dashed line on the figure) is readily apparent.

Finally, Table 4.18 gives the affine analysis indices for each test in the NSE series. The small expansion factors and fit errors indicate that the viewpoint independent reference frame is spatially stable in all the tests. The small average rotation values indicate that the reference frame alignment is correct.

Test	Rotation/mrd	ε/%	λ_{min}/cm^2	λ_{max}/cm^2
nse1	-0.233	-0.01	0.196	3.338
nse2	-0.228	-0.01	0.402	3.428
nse3	-0.224	-0.01	0.283	3.524
nse4	-0.219	-0.01	0.348	3.602
nse5	-0.214	-0.01	0.241	3.598

Table 4.18. Affine Analysis Indices for the NSE Tests.

Conclusions for the NSE Tests.

The five NSE tests illustrate the motion resolution system's performance in the presence of random observer motion perturbation. In all the tests the system performs well, achieving a considerable position variance reduction for the observer and a moderate noise reduction (at worst eight times, with respect to the measurement noise) for target positions. The velocity extraction performance of the system was good; and the temporal and spatial stability of the viewpoint independent reference frame was satisfactory.

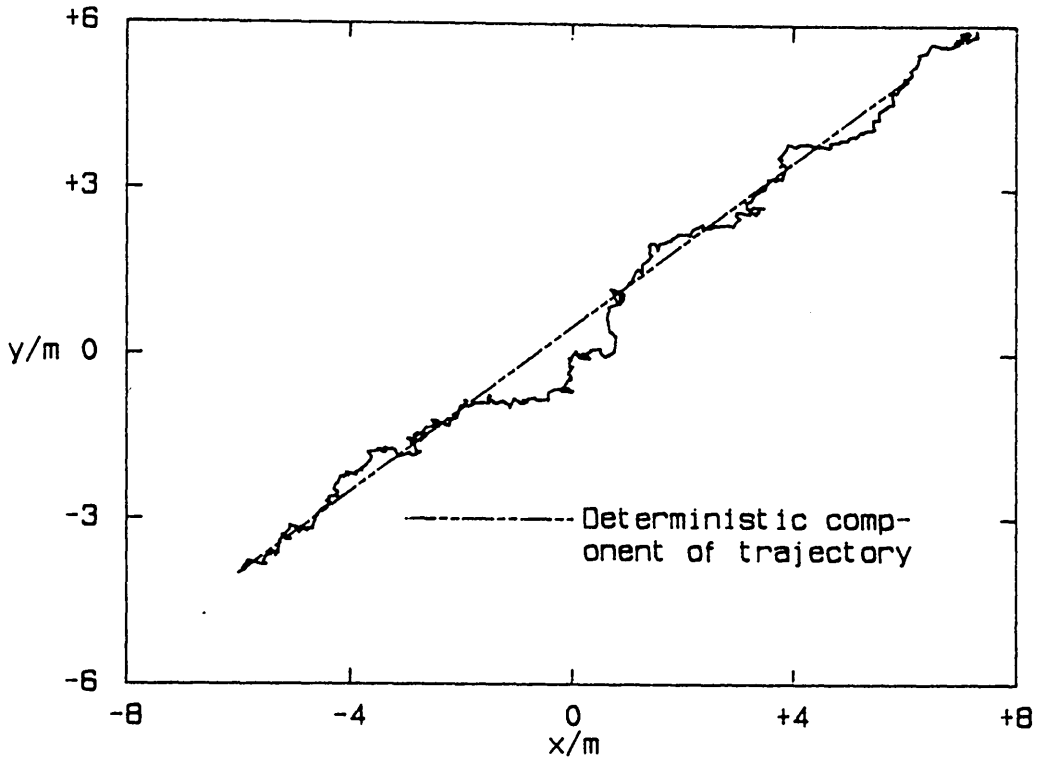


Figure 4.10. The Observer Trajectory in Test nse5.

An interesting anomaly in the set is that the velocity errors of those targets towards which the observer was moving decreased as the perturbation of the observer trajectory increased. The velocity errors for the remaining targets increased with increasing observer trajectory perturbation as expected.

4.5 An Assessment of the System's Performance.

The experiments described in the previous section were done with the twin objectives of verifying the theoretical results of section 4.1 and of assessing the performance of the new implementation of algorithm A described in section 4.3. This penultimate section examines the conclusions of the individual experiments and relates them to the theoretical discussion of algorithm A. Some areas requiring further study are also noted.

4.5.1 Theory and Practice.

The main conclusion to be drawn from the results presented is that the motion resolving algorithm is able to establish and maintain a spatially and temporally stable absolute reference frame provided that a number of stationary targets are correctly identified. In that case, stationary targets, moving targets, and non-deterministic observer motion are all handled correctly and the implementation under test achieved a substantial reduction in both measurement noise (in all the experiments) and observer motion noise (in the NSE tests).

These results address the points raised in section 4.1.7 concerning the stability and noise reduction properties of the algorithm. The experiments presented confirm the conjecture that if no moving targets are misclassified and some stationary targets are correctly identified then the algorithm may be expected to converge to a stable viewpoint independent reference frame and show some noise reduction.

The principal deficiency exposed by the experiments is in the hypothesis testing mechanism. None of the tests failed because moving targets were wrongly classified as stationary, but a substantial proportion of the more complex tests in the ML series failed because all the stationary targets were judged to be moving by the hypothesis test. These test failures confirm the necessity of correctly classifying some stationary targets.

There are two further points to make concerning the performance of the system in the ML test series. First, the hypothesis test used in the implementation is a simple one. The average absolute velocity and the velocity estimation covariance for each target were estimated from the sequence of target absolute state values using a fading memory filter. The mean velocity was then compared with the estimation error covariance using a χ^2 test and if the mean was found to be significant the moving target hypothesis was accepted.

The motion resolving algorithm easily tolerates the misclassification of stationary targets (provided, of course, that some are correctly classified), as test ml13 indicates, but it is much less tolerant of errors in moving target classification since these force the algorithm into an initialisation phase. The hypothesis test must therefore be conservative with regard to misclassifying moving targets, at the expense of possibly misclassifying some stationary ones. A conservative ad hoc threshold value of 1.2 was found by experiment to be appropriate (the expected value of the χ^2 statistic is 2).

The evidence of the ML test series suggests that the hypothesis test is excessively conservative. However, the situation is not as simple as that, for the tests reported in Chapter five (where observer angular motion was permitted) suggested that the same hypothesis test was excessively liberal.

The second point concerning the ML tests is directly relevant to this last observation. It is that the poor performance of the hypothesis test is aggravated by the small total number of targets in the tests and by the high proportion of moving targets.

The simple conservative hypothesis test frequently misclassifies stationary targets for short periods because of the random fluctuation of the χ^2 hypothesis value. However, with only four stationary targets present, the loss of information entailed by such misclassifications is serious. Additionally, the large proportion of moving targets aggravates the velocity extraction errors during the initialisation phase and may result in all the stationary targets being misclassified when the algorithm enters normal operation.

In a practical system these effects would be less serious. Greater numbers of targets will improve the likelihood of several stationary targets being correctly identified. More sophisticated hypothesis tests may be used to overcome the deficiencies of the simple scheme employed here (some suggestions for improving the hypothesis testing are given in section 5.5.2) and a careful screening of the targets used for initialising the algorithm would reduce the initialisation phase velocity errors.

4.5.2 Topics Requiring Further Study.

A representative set of performance experiments have been presented to give a fair picture of the practical performance achieved by the motion resolution system. However, there are a number of areas where further study is necessary to complete the assessment of the system. The aspects of the system that have not been investigated yet include:

- non-deterministic target motion;
- complex observer trajectories;
- the relationship of output estimation errors to input measurement noise.

4.6 Summary.

In this chapter a solution to the linear motion resolution problem has been described informally and defined formally as a linear recursive filter. Monte Carlo simulation experiments have shown that the algorithm achieves its goal of establishing and maintaining a consistent and stable viewpoint independent reference frame provided that some stationary targets are correctly identified and no moving targets are misclassified as stationary. The major weakness of the implementation assessed in the chapter was in its simple hypothesis testing mechanism.

Chapter 5. Dealing with Rotary Observer Movement.

In Chapter four the motion resolution problem was discussed. There a two-dimensional solution was presented for point objects and an observer subject to the restriction that the observer orientation be known *a priori*. Using that solution as a starting point this chapter considers the unrestricted motion resolution problem for two dimensions. Point objects retain their two degrees of translational motion freedom and the observer acquires an extra degree of rotational motion freedom since its measurement coordinate system is now free to rotate with respect to the world (Figure 5.1).

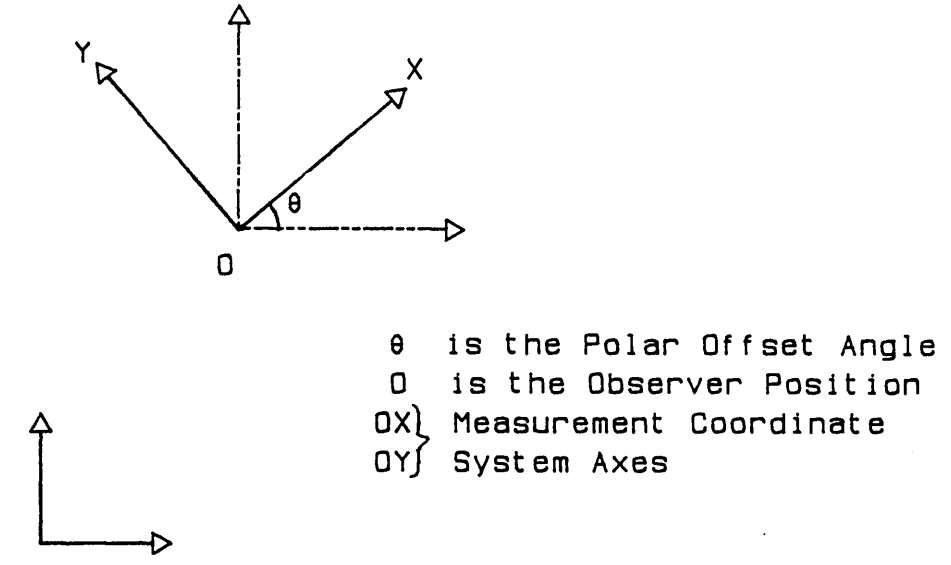


Figure 5.1. The Observer Measurement Coordinate System.

This chapter is divided into five sections. The first discusses the implications of admitting observer rotation. In the second and third sections I explore and evaluate two avenues of solution to the motion resolution problem. Section 5.4 presents experimental results of performance tests carried out on the motion resolution system developed in section 5.3, and the last section summarises and concludes the discussion of two-dimensional motion resolution in this dissertation.

In this chapter, as in Chapter four, the world in which the observer moves is assumed to be two-dimensional. However, it is mathematically convenient to treat the observer position and velocity as three-dimensional vectors. In what follows, therefore, these vectors are taken to have a zero z component and the observer and target motion takes place in the xy plane of a conventional, right-handed, Cartesian coordinate frame, the world reference frame. With this arrangement the observer angular velocity vector is parallel to the world frame z axis and its magnitude is equal to the rate of rotation of the observer. The vector points along the positive z axis for an observer turning anticlockwise. The angular velocity is completely determined by a single parameter, its third (z) component, and I shall therefore use vector and scalar symbols interchangeably for this quantity, according to the mathematical context.

5.1 The Effects and Implications of Observer Rotation.

The most obvious effect of permitting observer rotation is that the observer acquires an extra degree of freedom of motion. The orientation of the observer's measurement coordinate frame can and in general will vary continually as the observer moves. The motion resolution system must estimate and track the polar offset angle θ (Figure 5.1) that determines the relationships between measurement coordinates and the constant orientation viewpoint independent frame of reference.

The most important consequence of observer rotation, however, is non-linearity. Under the restriction that the observer orientation be known a priori, the motion resolution problem is both algebraically and dynamically linear. It is algebraically linear because the reflected motion field is linearly dependent on the observer velocity and the apparent motion field is linearly related to object velocity and the reflected motion field. It is dynamically linear because the observer rotation can be compensated using the a priori information, after which the expected target motions -- piecewise linear trajectories with random acceleration -- correspond to linear state transitions in the motion resolution system's tracking filters.

Observer rotation interferes with both these properties. It violates algebraic linearity since the reflected motion field comprises a non-linear rotational component. Dynamic linearity is violated because an observer rotation induces a circular apparent motion that corresponds to non-linear state transitions in the tracking filters. The consequences of these effects are explored in detail in the sections below.

5.1.1 Algebraic Non-Linearity.

Reflected motion was defined in Chapter three as that component of the apparent target motion which is caused by the proper motion of the observer. It is a three-dimensional vector field comprising a position independent translational part and a rotational part that is linearly dependent on position:

$$\underline{rm}(\underline{r}) = \underline{v}_0 + \underline{\omega}_0 \times \underline{r} \qquad (5.1)$$

where $\underline{v}_0$ and $\underline{\omega}_0$ are the observer linear and angular velocity vectors respectively. The observation equations, equations (4.1), are linear, so the focus of the algebraic non-linearity in the motion resolution problem is the vector cross product term, $\underline{\omega}_0 \times \underline{r}$, generating the position dependent rotational part of the reflected motion field.

The principal effect of this non-linearity is to complicate the inversion of the reflected motion field. The inversion problem is this:

> Given M noisy values of the reflected motion field, denoted by $\underline{rm}_j$ for j=1..M, obtained at estimated (noisy) positions $\underline{r}_j$ for j=1..M relative to the observer, estimate optimal values for the observer linear and angular velocity $\underline{v}_0$ and $\underline{\omega}_0$.

Using the standard least squares technique, the optimal values of $\underline{v}_0$ and $\underline{\omega}_0$ are those that afford a best fit reflected motion field, i.e. they minimise the total square error as a function of $\underline{v}_0$ and $\underline{\omega}_0$:

$$err(\underline{v}_0,\underline{\omega}_0) = \sum_{j=1}^{j=M} \left| \underline{rm}_j - \underline{rm}(\underline{r}_j) \right|_j^2, \quad |\underline{x}|_j^2 = \underline{x}^T W_j \underline{x}. \qquad (5.2)$$

The weight matrices W_j used in the vector normalisation depend on the errors expected in the residual vector $\underline{rm}_j - \underline{rm}(\underline{r}_j)$.

Unfortunately, the appropriate weight matrices in the expression (5.2) are themselves functions of the unknown angular velocity $\underline{\omega}_0$. The expected errors in the residuals are a combination of the noise in the reflected motion field estimate, $\underline{rm}_j$, and the errors in the corresponding position, $\underline{r}_j$; these latter errors are multiplied by components of $\underline{\omega}_0$ in the vector cross-product. Even in the two-dimensional situation where the direction of $\underline{\omega}_0$ is known *a priori* the estimation problem is a prohibitively complex and highly non-linear polynomial root computation.

Given the desirability of a weighted estimation that makes good use of the available error covariance information, there are two ways to proceed at this point: a suitable non-linear estimation technique may be used to estimate $\underline{v}_0$ and $\underline{\omega}_0$ directly from $err(\underline{v}_0,\underline{\omega}_0)$; or a heuristic solution technique may be employed. In the two-dimensional case the estimation of $\underline{\omega}_0$ reduces to the solution of a polynomial equation in the magnitude of $\underline{\omega}_0$ while in the three-dimensional case the problem comprises three coupled simultaneous polynomial equations in the components of $\underline{\omega}_0$. Suitable techniques exist for the solution of these types of problem (non-linear estimation techniques are reviewed by Bard (1974), for example) but they are computationally expensive iterative methods that are infeasible in view of the strong real-time constraint on interpreter processing.

The absence of fast analytic or numerical solution techniques makes the heuristic approach attractive, despite the sub-optimality that such a method entails. The heuristic adopted is to estimate $\underline{v}_0$ and $\underline{\omega}_0$ separately rather than jointly.

The estimation of $\underline{v}_0$, given a predicted value of the observer angular velocity, proceeds analogously to the solution presented in Chapter four for the case $\underline{\omega}_0=\underline{0}$. In that chapter, the reflected motion field contained only translational motion components and it

was inverted using the Information Averaging Filter represented by equation (4.2). The extension to this procedure to handle non-zero values of $\underline{\omega}_0$ is to remove the rotational components from the reflected motion field, using a correction term based on the predicted angular velocity, and then to apply the inversion equation (4.2) as before. If the predicted angular velocity is $\tilde{\underline{\omega}}_0$, (where, as in Chapter four, the tilde-caret convention indicates that this prediction is based on all the data received strictly before the current measurement set) then the correction term is a velocity $\tilde{\underline{\omega}}_0 \times \underline{r}_j$ which must be subtracted from the reflected motion velocity vector $\underline{rm}_j$. The estimate of $\underline{v}_0$ used by the system, denoted by $\hat{\underline{v}}_0$, is therefore given by

$$\hat{\underline{v}}_0 = \sum_{j=1}^{j=M} \beta_j \left(\underline{rm}_j - \tilde{\underline{\omega}}_0 \times \underline{r}_j\right) \quad \text{where} \quad \beta_j = \left| \sum_{j=1}^{j=M} P_j^{-1} \right|^{-1} P_j^{-1}, \qquad (5.3)$$

and the P_j are the bounds on the error covariances of the $(\underline{rm}_j - \tilde{\underline{\omega}}_0 \times \underline{r}_j)$. These P_j are calculated analogously to the covariance bounds described in section 4.1.6, but an extra term is included to account for the error in the correction velocity $\tilde{\underline{\omega}}_0 \times \underline{r}_j$. For simplicity in this calculation the errors in $\tilde{\underline{\omega}}_0$ and $\underline{r}_j$ are assumed to be independent.

A similar approach is taken for estimating $\underline{\omega}_0$ given a prediction of the observer linear velocity, say $\tilde{\underline{v}}_0$. The reflected motion field is compensated for observer linear motion and $\underline{\omega}_0$ is estimated from the resultant, $\underline{rm}_j - \tilde{\underline{v}}_0$. In this case, rather than tackle the non-linear polynomial estimation problem a second heuristic is employed -- a value for the observer angular velocity is calculated from each resultant field measurement using the identity (only valid in two dimensions)

$$\underline{\omega}_0 = \frac{\underline{r}_j \times (\underline{\omega}_0 \times \underline{r}_j)}{\underline{r}_j \cdot \underline{r}_j} \qquad (5.4)$$

and these computed values are combined using an Information Averaging Filter to give a composite estimate $\hat{\underline{\omega}}_0$ of observer angular velocity. Thus

$$\hat{\underline{\omega}}_O = \sum_{j=1}^{j=M} \beta_j \frac{\underline{r}_j \times (\underline{rm}_j - \tilde{\underline{v}}_O)}{\underline{r}_j \cdot \underline{r}_j} \quad \text{where} \quad \beta_j = \left| \sum_{j=1}^{j=M} \Omega_j^{-1} \right|^{-1} \Omega_j^{-1} \qquad (5.5)$$

and Ω_j is an upper bound on the error covariance of the j^{th} angular velocity value. The calculation of Ω_j can be simplified by neglecting the error contribution from the position vector $\underline{r}_j$. In practice, since the error covariance components are much smaller than the square of the range for most targets, this approximation is justified (for close targets a more accurate approximation could be used).

5.1.2 Dynamic Non-Linearity.

As the observer rotates, targets appear to move in trajectories containing circular components. The trajectories are in general cycloidal. The curved components in the trajectory are not catered for in the simple piecewise linear trajectory model used in Chapter four, and give rise to non-linear variations in the components of the

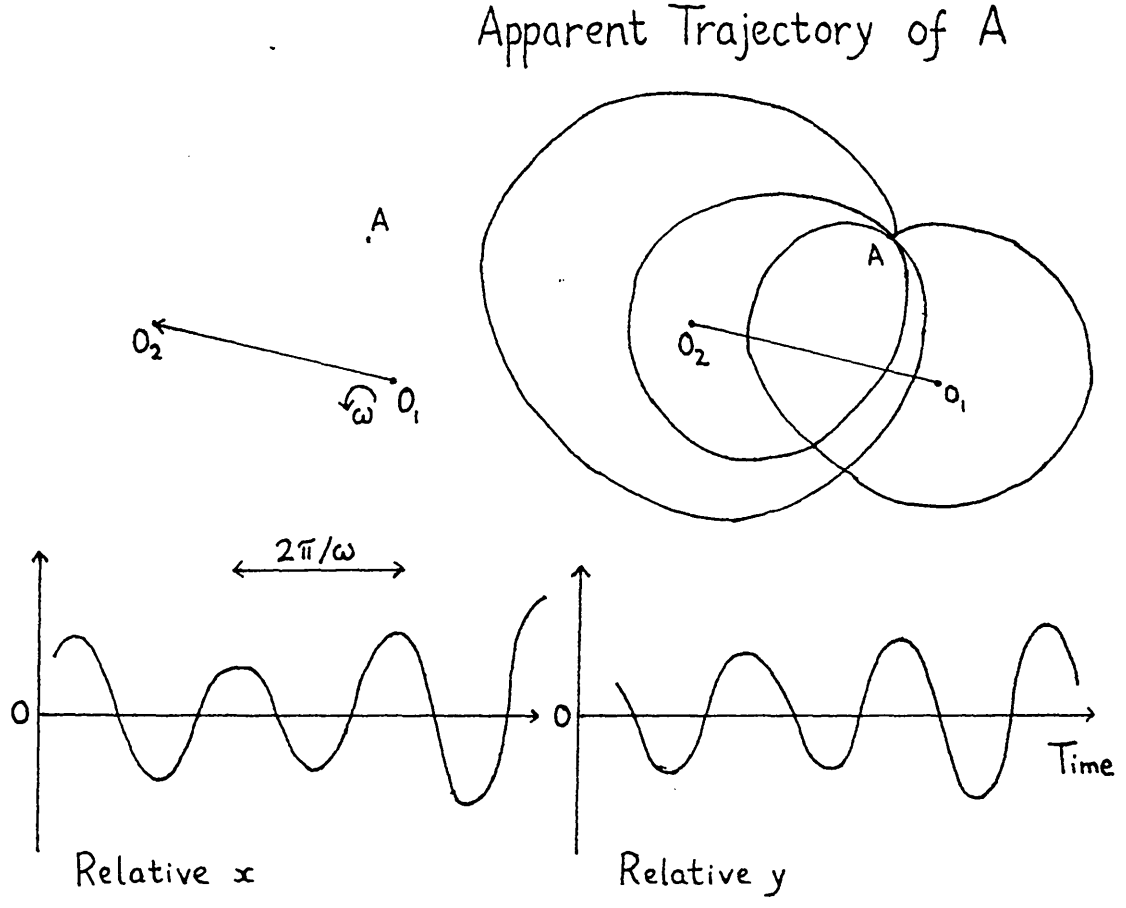


Figure 5.2. A Cycloidal Target Trajectory.

filter state vectors. For example, Figure 5.2 illustrates the apparent trajectory of target A as the rotating observer moves from $O_1 \rightarrow O_2$ and shows the variation of the relative position components of the target state. The linear and sinusoidal elements in the state variation are clearly visible. I have termed this induction of non-linear state variation dynamic non-linearity.

Dynamic non-linearity arises whenever a target trajectory cannot be described using a linear state transition model as required by the Kalman tracking filter. It causes biassed acceleration dependent components to appear in the filter innovation sequence and results in sub-optimal performance by the tracker. Effectively, the target is continually manoeuvring.

There are two basic ways to improve the situation: the state transition model may be altered to describe the target trajectory; or the tracking coordinate system may be changed to one in which the target trajectory corresponds to linear coordinate variations. Both techniques may be applied simultaneously.

The simplest modification to the state transition model to overcome the continual target manoeuvre caused by dynamic non-linearity is to ignore the error but increase the filter state transition noise to allow for the manoeuvre, possibly increasing the rate of measurement of the target position to reduce the prediction errors. A more sophisticated approach is to augment the filter state vector to include an estimate of track curvature and employ an Extended Kalman filter for tracking, with a linearised model of the curved trajectory. The simple approach suffers from biassed innovations, sub-optimal filter performance, and an artificially raised error covariance, but is adequate if the data rate is sufficient. The extended Kalman filter gives optimal tracking performance on the linearised trajectory at the expense of computational complexity.

The second method of modification is to choose an appropriate tracking coordinate system. In the linear motion resolution situation, the natural choice of coordinate system was a Cartesian frame common to all targets. Dynamic non-linearity arises when observer rotation is permitted because the cycloidal trajectories of targets

relative to the observer are curved and the target positions and velocities vary non-linearly as they traverse their trajectories. If only rotary motion were present, a polar coordinate system centred on the observer would be suitable -- the target trajectories would correspond to linear variations of radial and angular coordinates. However, the linear component in the trajectory gives rise to dynamic non-linearity in a polar tracking coordinate system, linear motion producing polar bias accelerations. Thus neither the common Cartesian nor the polar tracking frame is free from non-linear effects.

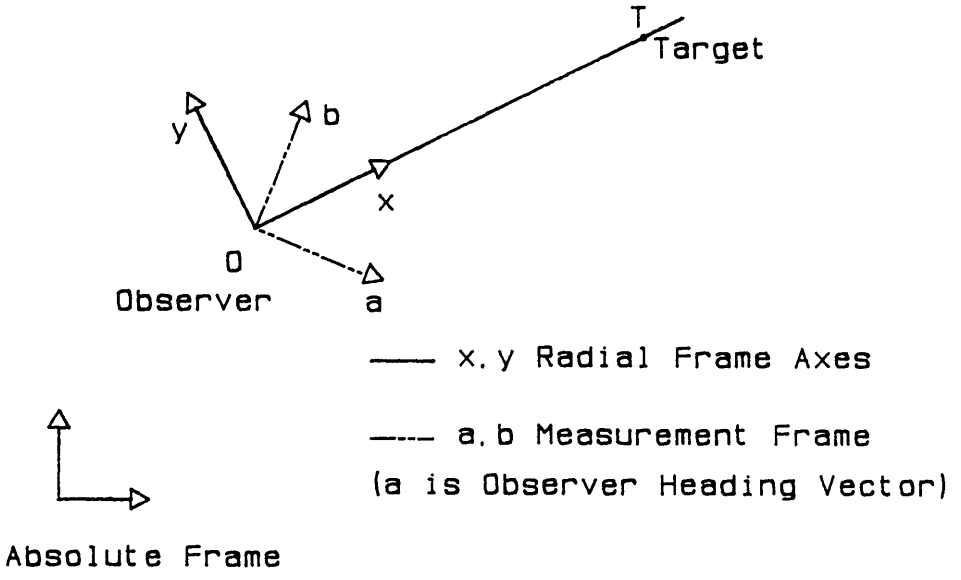


Figure 5.3. Target Radial Frame Alignment.

A third candidate for the tracking coordinate frame is illustrated in Figure 5.3. This is the target radial frame, a Cartesian frame with axes aligned along and across the radial direction to the target. For a stationary, non-rotating, observer the target radial frame combines the dynamic linearity of a Cartesian frame with the computational advantages of a polar frame (because measurement errors are uncorrelated between spatial dimensions in polar coordinates the tracking filter can be decoupled). It was used for this purpose by Holmes (1977) who went further, replacing the decoupled Kalman filters by simpler α-β tracking filters.

When the observer is allowed to move and rotate, however, the advantages of this frame disappear. Linear motion is hard to handle

in the radial frame; there is a question of the 'target direction' along which axis alignment occurs, since the target position is unknown (Holmes used the predicted target position); and coordinate transformations are required for each target whenever a prediction or estimation of the target state is done and whenever the observer motion estimates from targets are combined.

5.1.3 Summary.

In summary, observer rotation causes algebraic and dynamic non-linearity. The former necessitates the use of heuristic methods to obtain observer motion estimates from the reflected motion field. For the latter, a number of techniques have been considered, none of which is especially attractive for the situation where the observer possesses both linear and rotational motion. At this point, therefore, some simulation tests were carried out to determine the extent of the difficulties caused by both kinds of non-linearity.

5.2 An Experimental Investigation of Non-Linearity.

The linear motion resolving system described in Chapter four was taken as the basis for an experimental linear and angular motion resolving scheme. The enhanced system formed the testbed for an evaluation of the effects of algebraic and dynamic non-linearity on the performance of the motion resolving algorithm.

5.2.1 System Enhancement for Angular Motion Resolution.

The testbed system follows the architecture of the linear system (see section 4.3, page 93ff), partitioning processing into a kernel associated with observer tracking and channels that contain the necessary data and mechanisms for tracking a target. When a particular target is seen its channel is activated, and all active channels contribute information to the kernel and receive updated kernel estimates when observer track processing completes.

The changes made to the kernel are simple. An Information Averaging Filter is added, having the task of combining the angular velocity estimates supplied by active channels into a composite observer angular velocity estimate. The current observer angular velocity is integrated to provide an estimate of the observer's polar offset

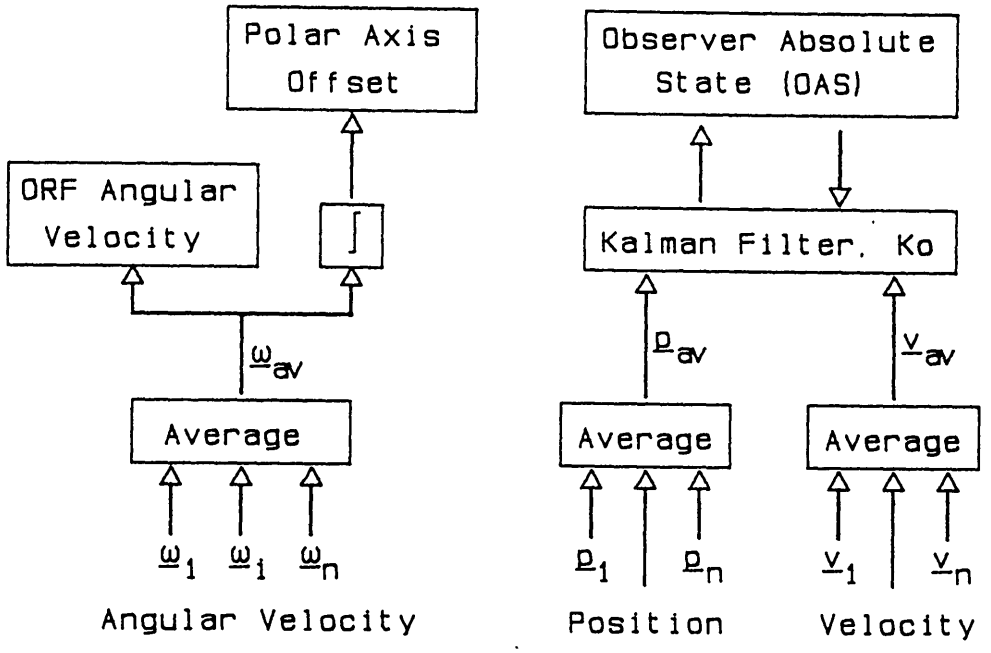


Figure 5.4. Block Diagram of Modified Kernel Processing.

angle. The block diagram of this processing is shown in Figure 5.4, where the linear kernel processing also appears.

Modifications to the target channels are more complex. The new channel structure is shown in Figure 5.5, from which it can be seen by comparison with Figure 4.5 that the relative tracking section of the channel remains unaltered. The target is tracked relative to the observer by Kalman filter Kr to provide an estimate of the apparent motion field. Unlike the linear system, however, the observer measurement coordinate system is now free to rotate as the observer orientation changes. I shall call this frame the observer-relative frame (the ORF).

The estimates passed from each channel to the kernel and the target and observer absolute state vectors are represented in the absolute coordinate frame. However, the apparent motion field estimates are in observer-relative coordinates, so the first step in computing estimates to pass to the kernel is to transform the target relative

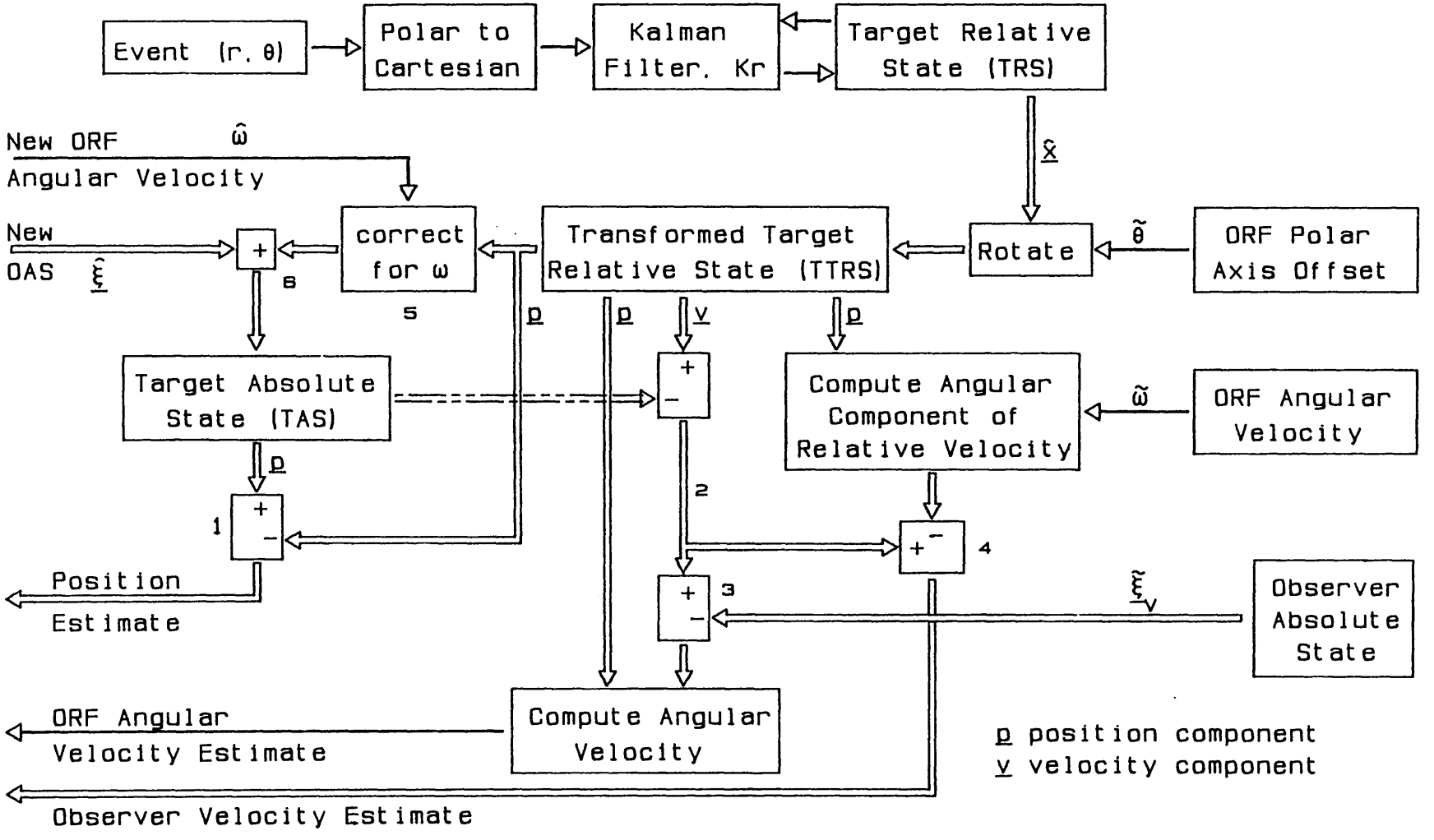


Figure 5.5. Augmented Channel Processing

state, the state vector of Kr, ($\hat{\underline{x}}$ in Figure 5.5) into the absolute frame. The current predicted observer polar offset angle $\tilde{\theta}$ is used to effect this transformation.

Once a transformed target relative state is available, three observer motion estimates are computed by each active channel.

1) The current target absolute position less the transformed target relative position is used as an observer position estimate. This is computed by box 1 in Figure 5.5.

2) The transformed relative velocity is compensated for target absolute motion at box 2 to give the reflected observer field, which is then corrected for the current observer linear motion in box 3 and used to estimate the observer's angular velocity (using the heuristic technique discussed above).

3) The reflected observer velocity from box 2 is corrected for the current observer angular velocity at box 4 and the resultant is used as an estimate of the observer linear velocity.

These three estimates are collated, over all active channels, by Information Averaging Filters in the kernel. Once kernel estimation is finished, the new observer state $\hat{\underline{\xi}}_0$ is added (box 6) to the transformed target relative states to give new target absolute position and velocity estimates, after a velocity correction for the observer rotation at box 5.

The effect of the varying orientation of the observer is depicted in Figure 5.6. Suppose that the orientation of the viewpoint independent frame (the absolute frame) is as shown and the observer is at the absolute frame origin. The relative position vectors for target A, as measured by the sensor attached to the observer, are illustrated for several observer orientations and it is clear that Kalman filter Kr will track the target in the observer-relative coordinate system. This means that the apparent motion field estimate held in the state vector of Kr is represented in a reference frame that rotates with the observer. Since the observer-relative frame is

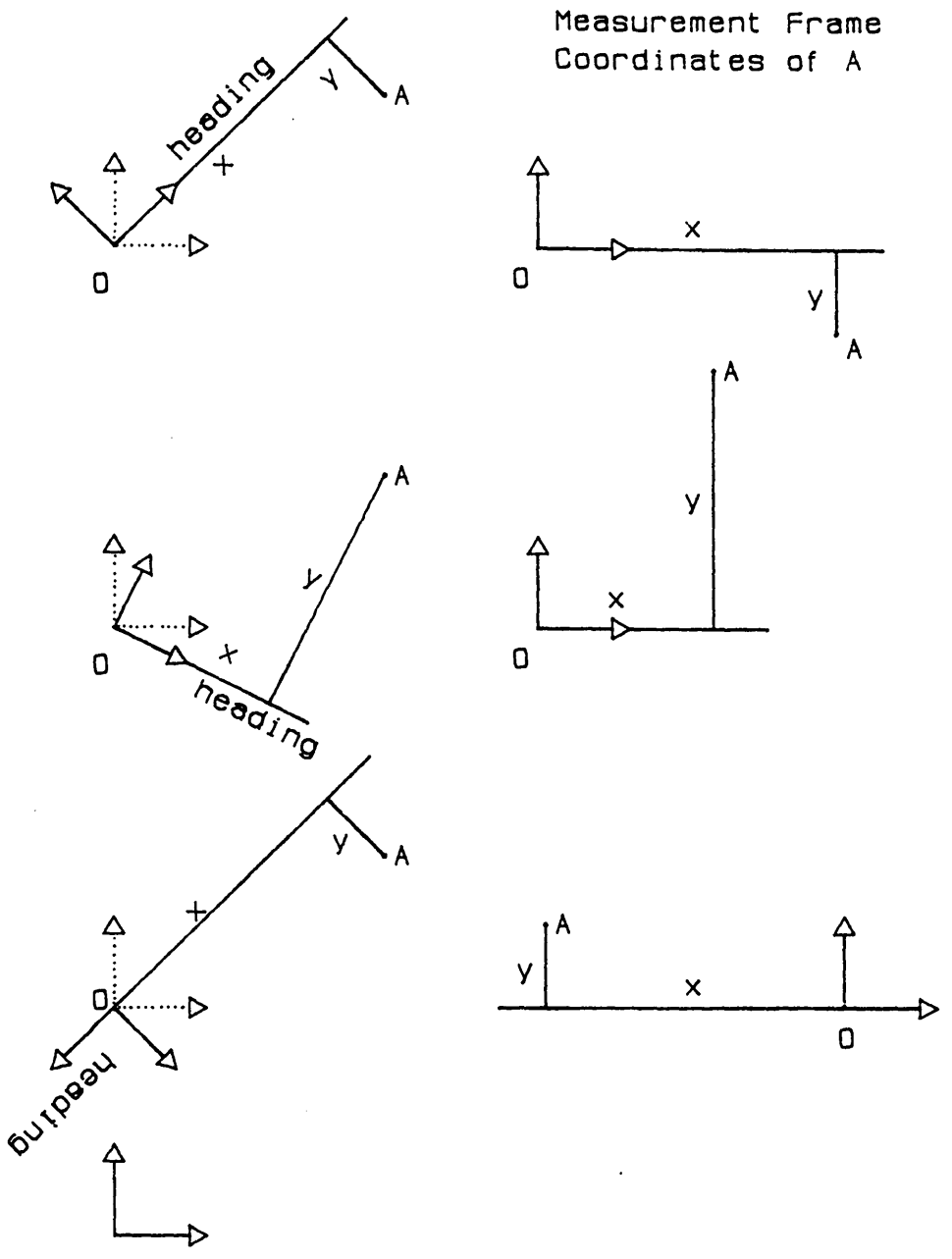


Figure 5.6 The Effect of Observer Rotation.

Cartesian, dynamic non-linearity effects may be expected when the observer rotates.

The angular velocity estimation in the channels differs slightly from the straightforward heuristic already described because some attempt is made to correct for dynamic non-linearity. The compensation scheme is illustrated in Figure 5.7. If the observer at O is rotating with angular velocity ω then the instantaneous velocity of

target A is $r\omega$ perpendicular to its position vector. Using the linear trajectory model employed by the Kalman filter Kr, and assuming that the filter state velocity agrees with the true velocity when A is in the position shown, the filter will predict the position of A, after time τ has elapsed, at P: $(r, r\omega\tau)$. However, the actual position of the target is $(r\cos\omega\tau, r\sin\omega\tau)$. The prediction error due to dynamic non-linearity in this situation is thus approximately

$$r\frac{(\omega\tau)^2}{2}\left(1, \frac{\omega\tau}{3}\right)$$

for small values of $\omega\tau$. The effect is thus proportional to the distance from the observer, and roughly proportional to the square of $\omega\tau$ for $\omega\tau << 1$.

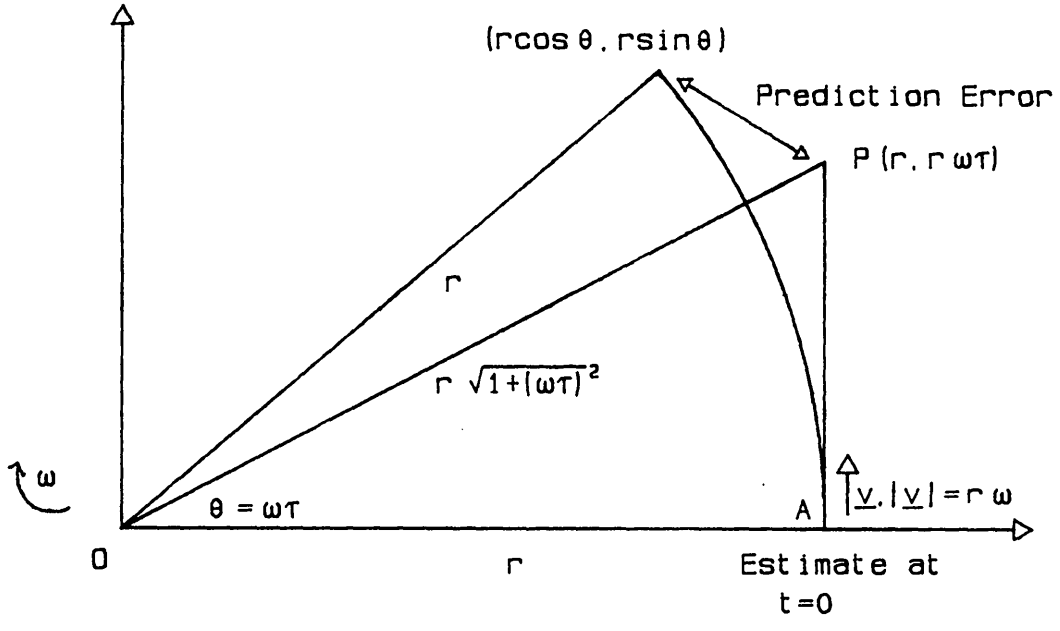


Figure 5.7. Dynamic Non-Linearity Compensation.

In actual operation the estimated position of the target lies between the true and predicted positions, so that the radial magnification is not as large as for the pure prediction; but the velocity is also deflected outwards with respect to its true value, so that the cross product in the angular velocity estimate is smaller than it ought to be. The net effect of dynamic non-linearity is to cause a shortfall in the angular velocity estimate which, for small values of

$\omega\tau$, is expected to vary roughly as $r(\omega\tau)^2$. This effect is compensated in the estimator actually used by the channels.

5.2.2 Tests and Results.

The new system was tested using similar data to the initial tests of the linear system. The test world comprised four targets set at the corners of a 70 metre square centred on the origin. The observer was positioned at the origin and remained there throughout the test. Its angular velocity was 50 milliradians per second (about 2.86° per second). The radial measurement error was 0.01 metres standard deviation (equivalent to ±3cm range resolution) and the angular error was 0.01 radians standard deviation (corresponding to a sonar beamwidth of about 3.4°). As in section 4.2.2, these parameters were chosen to simulate the performance of a medium resolution sonar. The test ran for 150 seconds of simulated time.

A graph of the estimated observer angular velocity, $\hat{\underline{\omega}}_O$, (the composite estimate using data from all four targets) against time is shown in Figure 5.8. Qualitatively, there is a short gap during which system initialisation occurs, then a settling down period, and the system remains in a steady state for the remainder of the test. Two conclusions may be drawn from the graph: the system appears to have settled down within 20 seconds of the start of the test, and the steady state angular velocity estimate is about 3.45% short. The former conclusion, taken with the 2.5Hz sonar rate, suggests that about 50 observations of the four targets are required for the system to stabilise. The latter conclusion implies that the system's estimated polar offset angle falls systematically behind the true polar offset angle as time progresses, and so the viewpoint independent reference frame rotates slowly with an angular velocity equal to the instantaneous shortfall in the current estimate.

The deficient angular velocity estimate is a serious problem -- it prevents the system from achieving a stable viewpoint independent

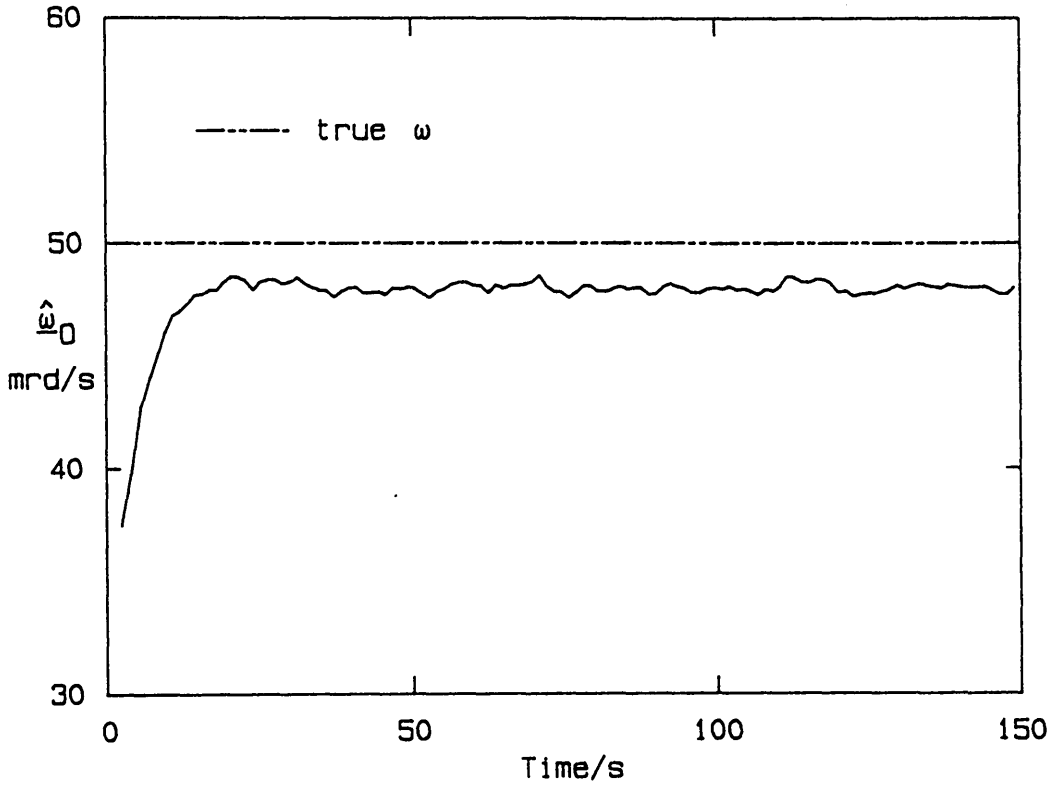


Figure 5.8. Variation of $\hat{\omega}_0$ with Time.

reference frame. To explore this difficulty further the dynamic non-linearity correction built into the estimator was removed, and a set of twelve tests run for varying angular velocities and world square sizes. The mean angular velocity estimated in each test, using values of the angular velocity estimate from t = 20 seconds to the end of the test, is tabulated in Table 5.1 against the true angular velocity, ω, and the size of the world square. Two effects are

	Angular Velocity/mrds^{-1}			
True, ω	-50	25	50	75
World Size	Estimated, $\hat{\omega}_0$			
40 m	-48.95	24.86	48.94	71.57
70 m	-48.01	24.74	48.01	68.69
100 m	-47.06	24.61	47.07	65.93

Table 5.1. Mean Estimated Angular Velocity in the Twelve Tests.

apparent in the table.

i) The shortfall in angular velocity increases as the world size increases. Plotting the deficit against world size (in Figure 5.9, for tests with positive ω) shows that the relationship is linear.

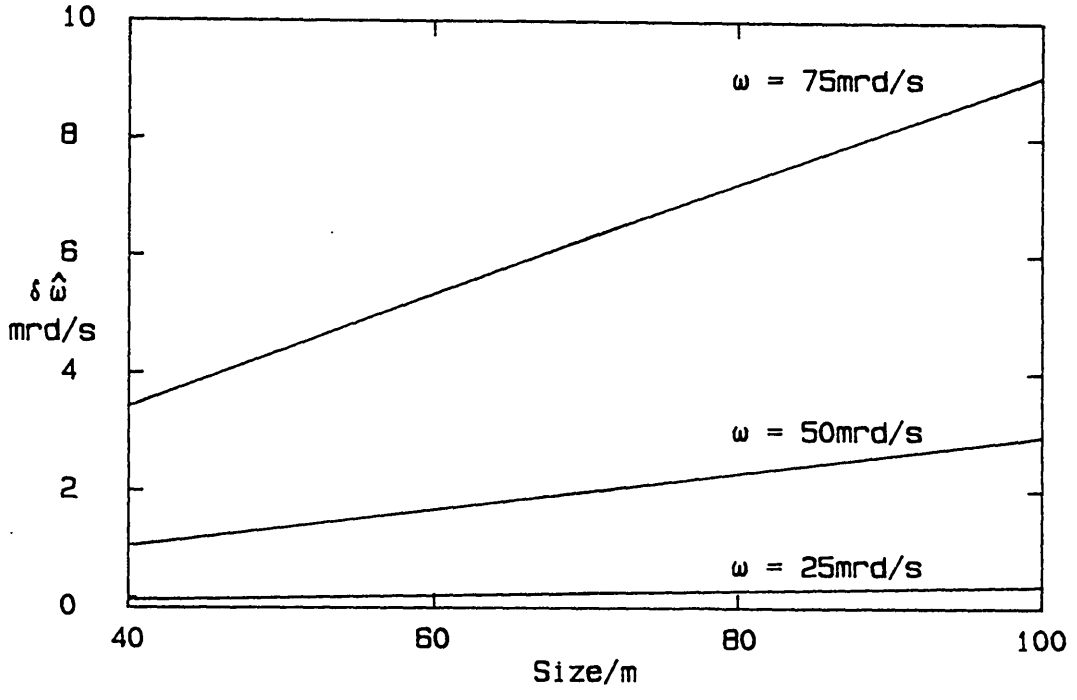


Figure 5.9. Angular Velocity Deficit vs. World Size.

ii) The deficit increases as the magnitude of the true angular rate increases. A log-log plot of deficit, $\delta\hat{\omega}$, over world size against the true angular rate shows that this relationship is a geometric one (Figure 5.10, also for tests with positive ω only).

The angular velocity deficit for the range of values tested may be described by

$$|\delta\hat{\omega}| = \beta r|\omega|^{\alpha} \qquad (5.6)$$

where r is the relative distance to the targets (equal to the world size over $\sqrt{2}$), and α and β are slope and intercept parameters obtained from the log-log graph. The best-fit values of α and β for

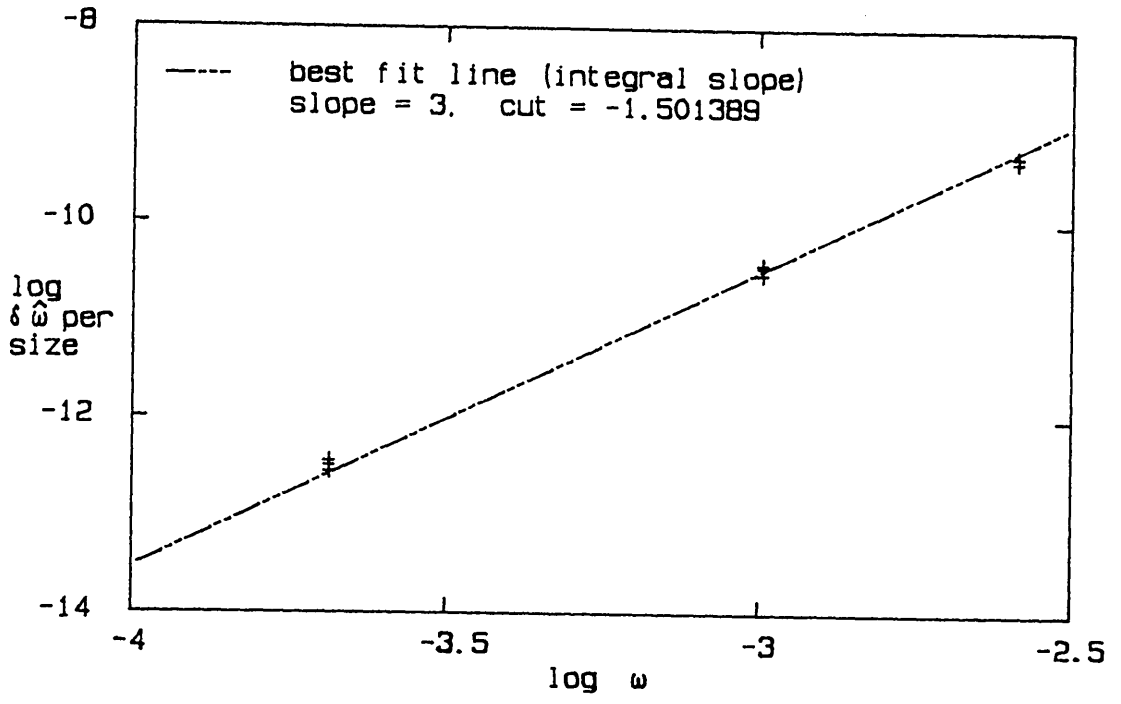


Figure 5.10. Log-log Graph of $\delta\hat{\omega}$/size vs. True Angular Velocity.

these tests are 2.904 and 0.2360 respectively. Restricting α to take integral values, the best-fit values are 3 and 0.3151; the corresponding line is shown dashed on the graph in Figure 5.10. In view of the unbounded right hand side of equation (5.6), that equation is not the true relationship between the deficit, world size and true angular velocity, but rather a local approximation valid for the range of values covered in the tests.

The good fit of the integral slope line suggests that the relationship described by equation (5.6) has a physical significance. In fact, the implication of (5.6) is that the relative deficit in the angular velocity estimate is proportional to the radial acceleration of the target, thus:

$$\left|\frac{\delta\hat{\omega}}{\omega}\right| = \beta r\omega^2.$$

This in turn suggests that the angular velocity shortfall is directly due to some dynamic non-linearity effect.

5.2.3 Theoretical Investigation of the Angular Estimator.

While equation (5.6) provides the information necessary to correct the angular velocity estimate in a range of situations, it does not explain the deficit: it merely describes it. There are two possible causes for the effect: dynamic non-linearity, causing large estimated position vectors and small cross-product terms in the estimator (5.4), or algebraic non-linearity which could bias the estimator. It is likely that both causes contribute, although in view of equation (5.6) the dominant effect appears to be the former, so a brief theoretical investigation of the latter possibility was undertaken. The mathematical details can be found in Appendix D.

The angular velocity estimator used in the system is

$$\hat{\omega} = \frac{\underline{r}\times\underline{v}}{\underline{r}.\underline{r}}$$

where $\underline{v}$ is a target velocity and $\underline{r}$ is the corresponding target position. This expression is linear in the components of $\underline{v}$ and non-linear in the components of $\underline{r}$. Both $\underline{r}$ and $\underline{v}$ contain noise components; if these are correlated they will interact in the vector cross-product.

As a first approximation, let the noise components in $\underline{v}$ and $\underline{r}$ be Gaussian and uncorrelated. Let the expected values of velocity and position be $\underline{v}_e$ and $\underline{r}_e$ respectively. Then the expected value of the angular velocity estimate is

$$\mathbf{E}\left[\ \hat{\omega}\ \right] = \mathbf{E}\left[\frac{\underline{r}\times\underline{v}}{\underline{r}.\underline{r}}\right] = \mathbf{E}\left[\frac{\underline{r}}{\underline{r}.\underline{r}}\right] \times \underline{v}_e,$$

since the estimator is linear in $\underline{v}$ and the noises are uncorrelated. The remaining expectation is a double integral over the possible values of $\underline{r}$

$$\mathbf{E}\left[\frac{\underline{r}}{\underline{r}.\underline{r}}\right] = \iint \frac{\underline{r}}{\underline{r}.\underline{r}}\ \mathrm{pdf}(\underline{r})\ \mathrm{dS}$$

where $\mathrm{pdf}(\underline{r})$ is the (two-dimensional) Gaussian probability density

function of $\underline{r}$. This integral is evaluated in Appendix D, and the resulting expression for $E\left[\hat{\omega}\right]$ is:

$$E\left[\hat{\omega}\right] = (1-e^{-M})\frac{\underline{r}_e \times \underline{v}_e}{\underline{r}_e \cdot \underline{r}_e}, \quad \text{where} \quad M = \frac{\underline{r}_e \cdot \underline{r}_e}{2\sigma^2} \tag{5.7}$$

and σ^2 is the position noise variance.

Thus, under the simplifying assumptions of uncorrelated Gaussian position and velocity noise the estimated angular velocity is given by equation (5.7). The estimate is biassed by a term $\hat{\omega}e^{-M}$. Relaxing the assumption of uncorrelated noise makes the integral significantly more complex but the analysis in Appendix D shows that the bias caused by noise correlation is proportional to $\frac{1}{M}$.

Now in the 70 metre square world, for typical position noise variances estimated by the tracking filters (about 0.2 m^2), the exponent M is about 6000, and the exponential is effectively zero. Even for correlated position and velocity noise, the expected deficit is of the order of 0.1%, which is still much smaller than the observed shortfall.

The conclusion of the investigation is that algebraic non-linearity causes a bias in the angular velocity estimator only for large position noises and close targets. For the conditions obtaining in the tests, the bias terms from algebraic non-linearity are negligible. The deficit observed in the tests must therefore be due entirely to dynamic non-linearity effects.

5.2.4 Conclusions.

The experimental and theoretical investigations described above support several conclusions.

1) The effects of the dynamic non-linearity caused by the rotating observer-relative reference frame are severe and far reaching, preventing the motion resolving system from achieving a stable viewpoint independent reference frame. Effectively the observed targets are given a common rotational motion matching

the deficiency in the observer angular motion estimate.

2) The system nevertheless appears to stabilise its angular velocity estimate within the first twenty seconds of a test run.

3) Algebraic non-linearity plays no significant role in the angular velocity estimation bias seen in the tests.

4) The relative deficit in angular velocity appears to be proportional to the radial acceleration of the targets. This relationship could be used to correct the estimator for small angular velocities.

5.3 Compensating for Dynamic Non-Linearity.

The experiments highlighted the serious problems caused by a rotating observer-relative tracking frame. For rotation rates of 20 milliradians per inter-sample time interval (e.g. 50 milliradians per second and 0.4 seconds sampling interval) the observer angular velocity estimates computed by the system are seriously in error. This section explores a solution to this difficulty.

5.3.1 A Rotation-Compensated Tracking Frame.

In the experimental motion resolving system described above, the rotating tracking coordinate frame was the cause of the dynamic non-linearity effects observed. Target trajectories induced by tracking frame rotation result in non-linear state transitions and bias accelerations that disturb the operation of the tracking filters.

The underlying problem is that the tracking frame has been identified with the observer-relative measurement frame. The orientation of the latter is determined by the current observer orientation and varies as the observer rotates. The solution to the dynamic non-linearity difficulties is to break this identification by

transforming measurements into a rotation-compensated tracking frame using the current estimate of the observer's polar offset angle.

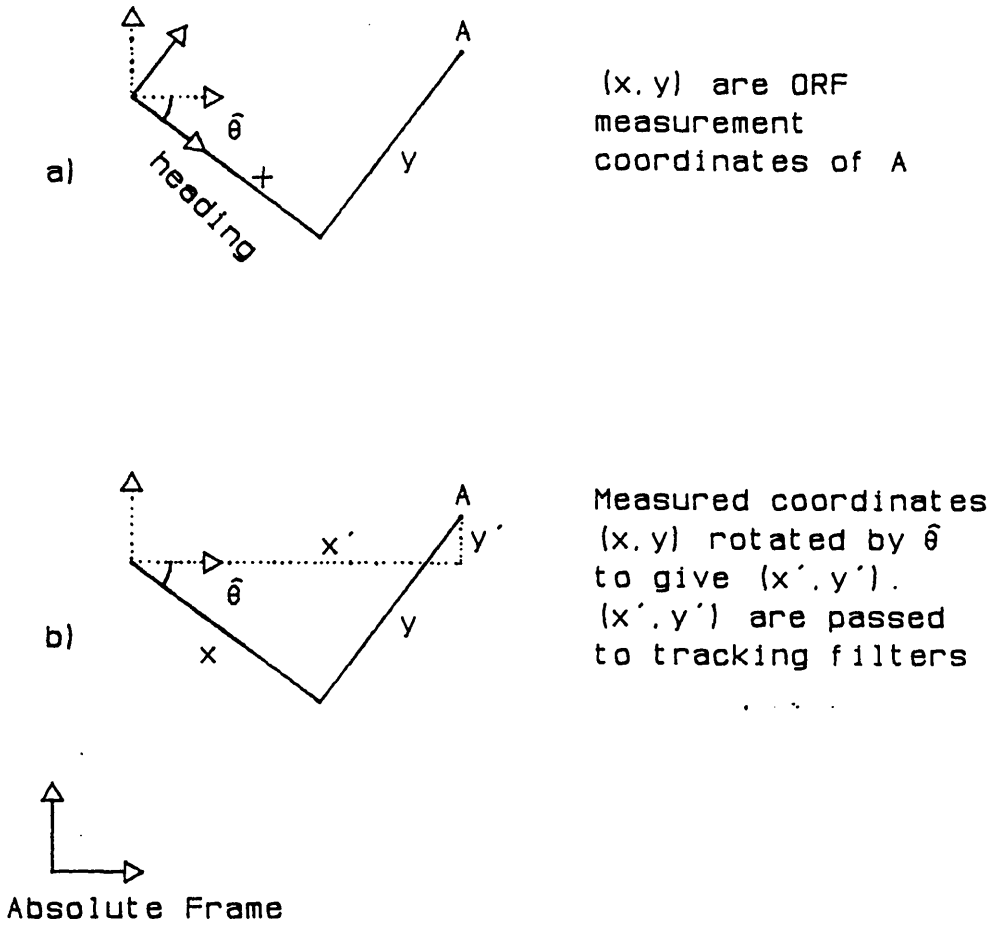


Figure 5.11. The Rotation-Compensation Transformation.

This process is illustrated in Figure 5.11. The observer at O measures the relative position of target A in Figure 5.11a and obtains the relative position coordinates (x,y) in the current observer-relative measurement frame. Since the ORF x axis is parallel to the observer heading vector, the measured relative position is the true relative position rotated by the true polar offset angle θ. The measured relative position is therefore rotated back using a current prediction $\tilde{\theta}$ of the polar offset angle, as shown in Figure 5.11b, and these transformed coordinates (x',y') are used by the relative tracking filter for target A.

When the relative motion tracking filters are fed with rotation-compensated measurements the apparent motion field they estimate contains only translational reflected observer motion. The non-linear rotational component of reflected motion is eliminated by the varying transformation from measurement to tracking coordinate system. If the polar offset angle is accurately estimated, then dynamic non-linearity disappears and the situation is effectively that of Chapter four.

In practice, this view of the situation is optimistic. The motion resolution problem in the rotation-compensated frame is only approximately linear. When polar offset estimation is perfect the tracking coordinate system is fixed in orientation (but not in position) relative to the viewpoint independent frame; when noise is present in the polar offset estimate, the tracking coordinate system orientation fluctuates about its nominal value. This randomly fluctuating misalignment of the tracking frame amounts to an uncompensated differential rotation, and it induces differential rotational components of reflected motion in the tracking filter states. (Alternatively, some observer reflected motion remains uncorrected because of the varying polar offset error.) The angular velocity corresponding to the differential reflected motion is the difference between the true and estimated observer angular velocities.

If the polar offset estimate is a good one, the differential angular velocity will be small and the induced differential reflected motion will also be small. The resulting dynamic non-linearity can be compensated by suitably increasing the tracking filter state transition noise. Further, the small differential angular velocity can be estimated accurately using the heuristic technique of section 5.2.1 and a correction for the small dynamic non-linearity based on equation (5.6).

The estimated differential angular velocity can be integrated to provide the polar offset estimate. A first step of integration converts the differential angular velocity to the observer angular velocity estimate; a second generates the polar offset angle.

5.3.2 Implementing Rotation-Compensation.

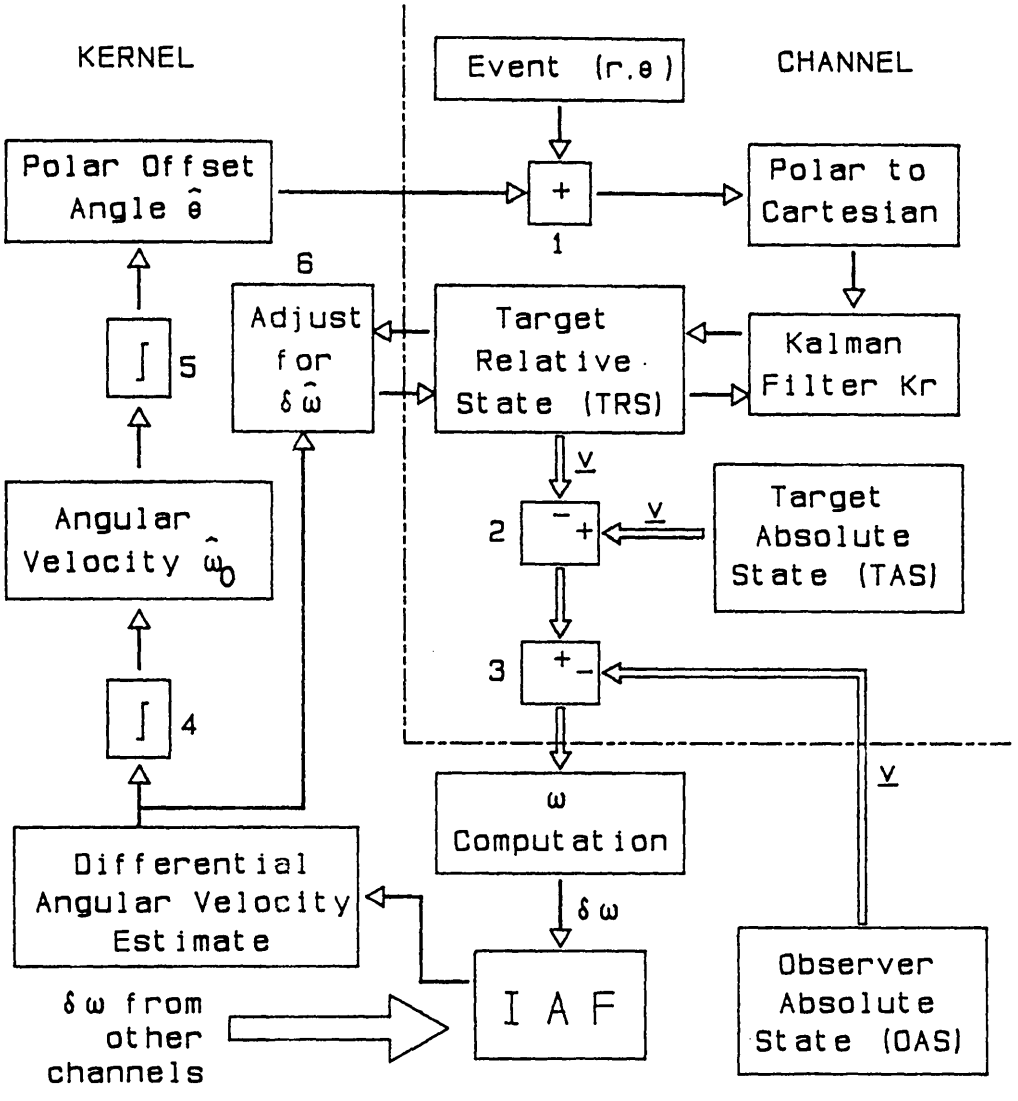


Figure 5.12. Rotation-Compensation Processing.

The processing required to achieve rotation-compensation is illustrated schematically in Figure 5.12. The linear velocity and position estimation parts of the system are omitted for clarity and the division between kernel and channel is marked by the dashed boundary.

The transformation from observer-relative measurement frame to rotation-compensated tracking coordinates is accomplished by adding the current polar offset prediction $\tilde{\theta}$ to each measured azimuth angle before converting the measurement from polar to Cartesian coordinates. This is done by box 1 in Figure 5.12. The error variance of the polar offset is treated as an additional angular measurement

error variance.

The tracking filter Kr operates in the rotation-compensated tracking frame which is nominally aligned with the viewpoint independent reference frame. The viewpoint independent, or absolute, frame is just the rotation-compensated tracking frame after the linear motion compensation achieved by the motion resolution system. The target relative state tracked by Kr is thus nominally recorded in the absolute frame and no transformation is needed to move estimates to that frame before passing to the kernel.

When a channel is activated it computes a differential angular velocity estimate $\delta\hat{\omega}$ using the target relative state velocity components, corrected for target and observer linear absolute motion at boxes 2 and 3 respectively. The kernel combines estimates from all currently active channels to give a composite differential angular velocity, and the current observer angular velocity estimate, $\hat{\omega}$, is adjusted by that amount using the integrator box 4. The polar offset angle is also adjusted so that the integration rule in box 5 is effectively a trapezium integration (a note is kept of the last time the observer angular velocity was estimated) thus:

$$\hat{\omega} = \tilde{\omega} + \delta\hat{\omega} \tag{5.8}$$

$$\hat{\theta} = \tilde{\theta} + \frac{\delta\hat{\omega}\tau}{2} \tag{5.9}$$

where τ is the interval between estimations. In between estimations of $\hat{\omega}$ the polar offset and angular velocity are predicted using the equations

$$\tilde{\theta} = \hat{\theta} + \hat{\omega}t \qquad \tilde{\omega} = \hat{\omega} \tag{5.10}$$

where t is the elapsed time since the last estimation.

Once the composite estimate of differential angular velocity is known the relative tracking filter states must be adjusted (done by box 6 in Figure 5.12). The reason for this is as follows. The velocity components of the target relative state vectors contain an element of differential reflected motion described by the new composite estimate $\delta\hat{\omega}$. Once the observer angular velocity estimate has been

updated using equation (5.8) this motion will be accounted for twice: once in the target relative states and once in the polar offset computation. The differential reflected motion induced by $\delta\hat{\omega}$ must therefore be subtracted from the target relative states. The result is that the rotation is represented only in the observer angular velocity. Thus the rotation-compensation process in 6 can be seen as a mechanism for transferring the non-linear element of the apparent motion field to the observer polar offset estimation where it can be treated linearly.

5.3.3 Initialising the Rotation-Compensation Loop.

The processing that achieves rotation compensation depends upon estimates of two quantities: the polar offset angle and the observer angular velocity. For simplicity these are initialised to zero. The initial polar offset angle determines the azimuth reference direction of the system and is thus arbitrary unless the viewpoint independent reference frame must register with an *a priori* map. The observer angular velocity may be set to zero because the loop will converge to a suitable value. This may be seen from the following argument.

Suppose that the angular velocity is not changing rapidly in relation to the sample rate of the sonar. Initially the system performs like the experimental system described in section 5.2, because the differential angular velocity is large, and estimates a differential angular velocity which is $1-\rho$ of the true angular velocity, for some fraction ρ less than unity. The remaining angular velocity to be transferred from the target relative states is ρ times the initial angular velocity. Because this differential is smaller than the initial value, and the angular velocity estimator is therefore less biassed, the second differential angular velocity estimate accounts for more than $1-\rho$ of the actual differential. At the third estimation, the outstanding differential angular velocity is less than ρ^2 of the initial value. Continuing this line of reasoning, the differential angular velocity to be estimated at sample $n+1$ is smaller than ρ^n of the initial angular velocity. The effect of the unknown initial angular velocity therefore dies away faster than the

geometric progression $\{\rho^n\}$ decays.

Once the rotation-compensation loop is initialised it exhibits a negative feedback behaviour. If the current observer angular velocity is low, as in the initialisation example analysed above, the differential angular velocity induces a forward rotation component in the target states and the positive estimated differential angular velocity causes the observer angular velocity to rise. When the observer angular velocity exceeds the true angular velocity a reverse induced rotation of the tracking frame generates a negative differential angular velocity to correct the error.

5.3.4 The Complete Motion Resolution System.

The schematic structure of the rotation-compensated motion resolving system is set out in Figure 5.13. The partition between kernel and channel processing is shown as a dashed line and the interaction of the motion resolution system with segmentation processing[1] is explicitly shown. The system shown in Figure 5.13 is complete: it handles both linear and rotary observer motion (for the two-dimensional world).

System operation is identical to that of the experimental system described earlier except for two important differences. The first is the rotation-compensation loop discussed above. The second is that the estimates of observer linear velocity computed by the channels are not corrected for the differential angular velocity -- this corresponds to an assumption that the differential angular velocity (and hence the dynamic non-linearity) is small. If this assumption is invalid in a particular situation, a more complex heuristic estimation of linear and differential angular velocity may be used. For example, the differential angular velocity estimate may be used to pre-correct the reflected motion field for the linear velocity estimation and if appropriate the correction and estimation cycle could be iterated a few times.

1. Possible realisations of this module are the subject of Chapter seven.

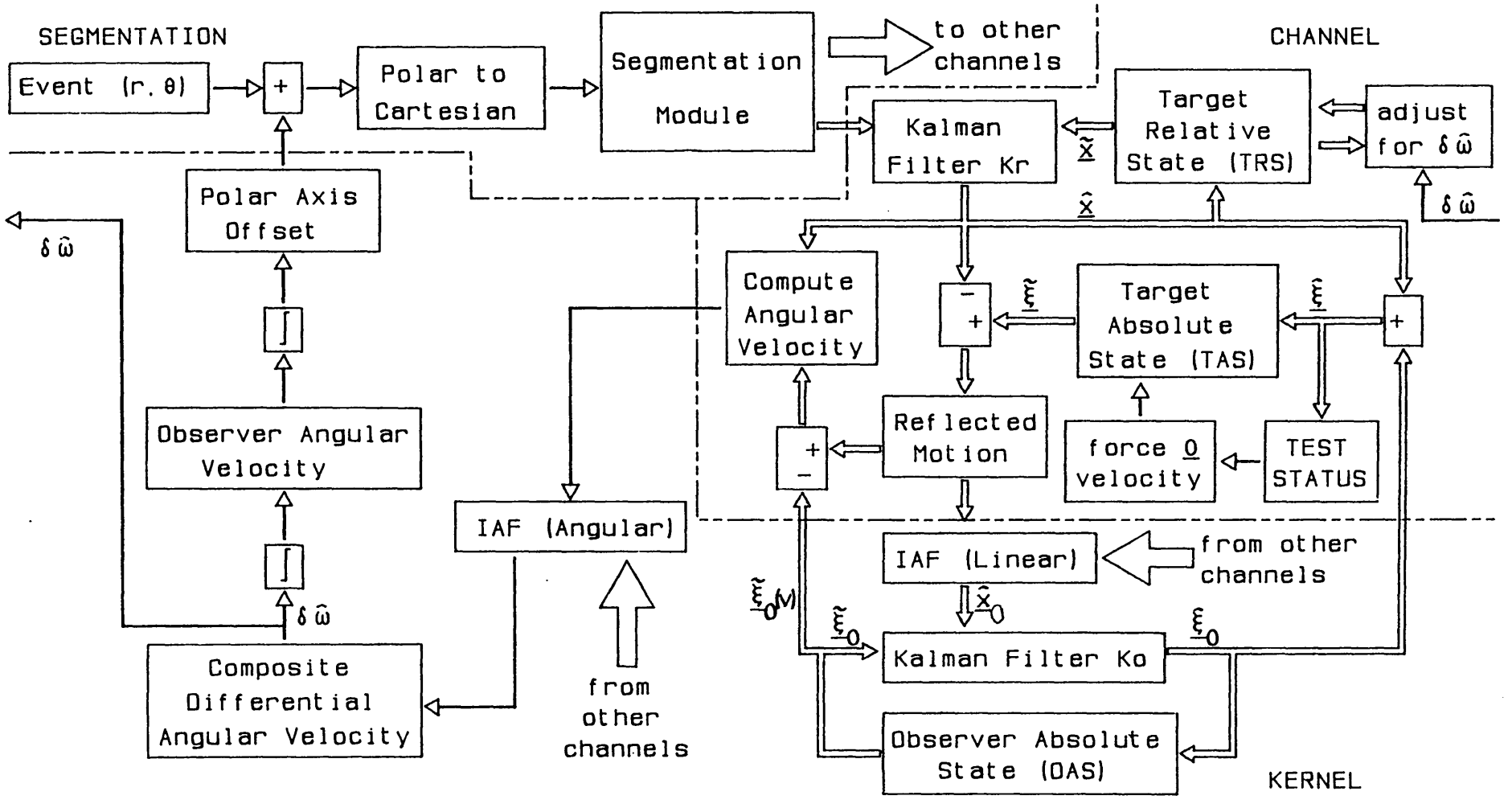


Figure 5.13. Schematic Diagram of the Complete Motion Resolution System

5.4 Performance Tests on the Complete Motion Resolution System.

The complete motion resolution system, incorporating rotation-compensation, was extensively tested by computer simulation. Five series of tests were done and the system performance was assessed in terms of the same categories as the linear system of Chapter four: velocity extraction, in this case including both linear and angular velocity; the reduction in measurement noise achieved; the temporal and spatial stability of the viewpoint independent reference frame established by the system; and the initialisation and settling time of the system.

Each of the five sections following describes one series of tests, presents the results obtained, and discusses the particular aspects of the system performance highlighted in those results.

5.4.1 Angular Velocity Extraction.

The first series of tests applied to the system was the set of angular velocity estimation tests used in section 5.2.2 (page 142) to evaluate the previous experimental system. In these tests the observer was positioned at the world origin with four objects in a square centred on the origin. Various observer angular velocities and world square sizes were used. The results of the fifteen tests in this series (nine of the tests from section 5.2.2 and six new related ones) are summarised in Tables 5.2 and 5.3; the angular velocity estimate from a member of the test set is shown graphically against time in Figure 5.14.

Qualitatively the graph is very similar to the experimental system graph shown in Figure 5.8. After six observations the target channels are initialised and begin to contribute estimates to the kernel (at $t = 2.5$ seconds) and there is a brief disturbance while the linear velocity estimate stabilises. The angular velocity estimate was initialised to the true value to minimise the system settling time and obtain the maximum of useful data from each test. From about $t = 20$ seconds the angular velocity estimate fluctuates

randomly about the true value.

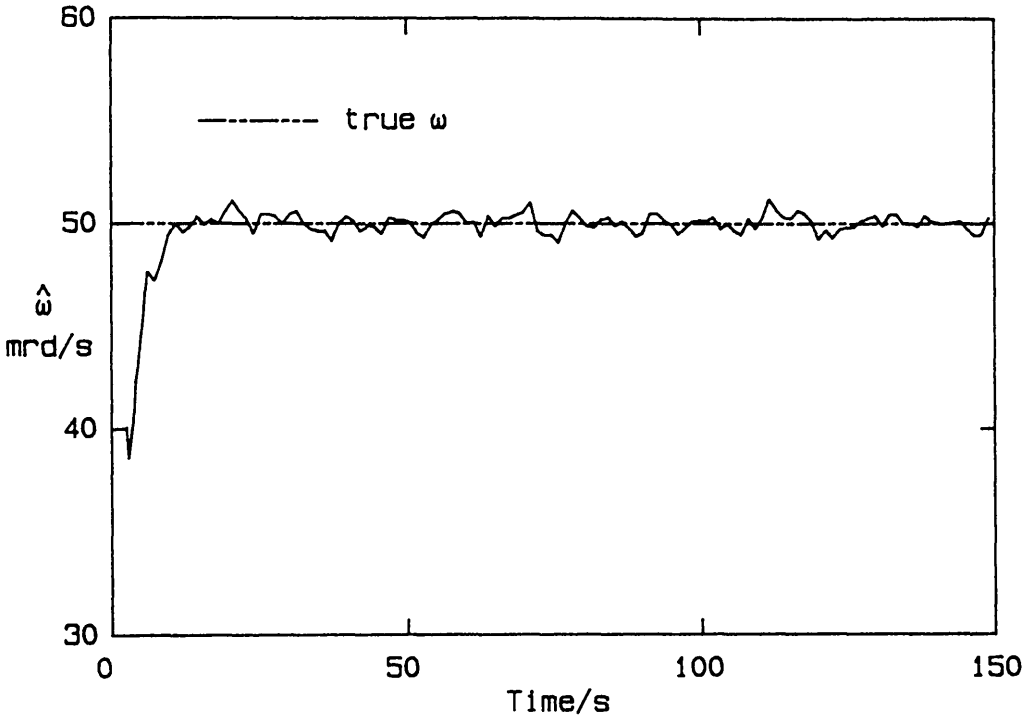


Figure 5.14. Variation of $\hat{\omega}$ with Time.

ω/mrds^{-1}	Test	World Size/m	$\hat{\omega}$/mrds^{-1}	Error Variance of $\hat{\omega}$/mrd^{2}s^{-2}	$\hat{\theta}$/mrd^2
	11w1	20	24.999	1.147	2.213
	12w1	40	25.003	0.405	1.532
25	13w1	70	25.006	0.176	1.176
	14w1	100	25.009	0.107	0.999
	15w1	140	25.011	0.067	0.866
	11w2	20	49.999	1.148	2.172
	12w2	40	50.003	0.404	1.527
50	13w2	70	50.007	0.177	1.142
	14w2	100	50.011	0.107	0.965
	15w2	140	50.015	0.068	0.842
	11w3	20	74.999	1.147	2.060
	12w3	40	75.003	0.404	1.553
75	13w3	70	75.007	0.178	1.121
	14w3	100	75.012	0.114	0.991
	15w3	140	75.020	0.076	0.738

Table 5.2. Angular Estimate Error Analyses.

Table 5.2 presents the results of an analysis of polar offset and angular velocity estimation error. For each test the mean value of $\hat{\omega}$ and the error variances of $\hat{\omega}$ and $\hat{\theta}$ are tabulated against the world size and the true angular velocity. The mean values highlight the long-term accuracy of the system in tracking the observer rotary movement; the variances describe the system stability and noise reduction. The noise in these tests comes from the 10 mrd standard deviation angular measurement error which, with the 2.5 Hz sampling rate, gives angular velocity errors of about $25\sqrt{2}$ mrds^{-1} standard

	Mean $\delta\underline{v}$/mms^{-1}		PEV/m^2		Position Offset/m	
Target	x	y	x	y	x	y
Obs	-0.085	0.148	0.00020	0.00025	0.019	0.004
1	-0.259	-0.336	0.015	0.025	-4.806	-5.599
2	1.150	-0.298	0.025	0.015	5.608	-4.810
3	-0.833	-0.556	0.019	0.031	4.830	5.589
4	0.586	-0.576	0.016	0.011	-5.571	4.819

Affine Transformation Analysis Results.

Mean $\delta\hat{\theta}$/mrd	Rotation/mrd	ε/%	λ_{min}/mm^2	λ_{max}/mm^2
151.97	149.23	0.01	0	31.17

Table 5.3. Linear Motion Error Analysis for Test 13w2.

deviation. It is clear from the results in Table 5.2 that the new system estimates the observer angular velocity correctly in all the tests.

In Table 5.3 the error analysis of object and observer linear motion is given for test 13w2. In these angular tests both objects and observer are stationary -- the mean velocities and positions illustrate the accuracy of estimation, while the variances describe stability and noise reduction properties as before. For this test the measurement noise standard deviation was about 0.50 metres in the angular direction, caused by the azimuth noise. The position offset vectors (the difference between the position estimated by the motion resolving system and the position known to the simulator) are large, but as the affine analysis shows they are caused by a rigid rotation

of roughly 0.15 radians.

The affine analysis indices also show the mean polar offset estimation error $\delta\hat{\theta}$ for comparison with the average rotation. The two values differ by 2.74 milliradians, which is therefore the amount (approximately) of the drift in the polar offset angle estimate during the settling time of the system. This drift in the viewpoint independent frame orientation, like the drift in its origin position, is arbitrary and occurs during the initialisation and settling period when the motion resolving system is attempting to establish the reference frame. Once the absolute reference frame is established, however, its origin position and the frame orientation are fixed, as the small error variances show. The rotation of 150 mrd defines the relative orientation of the viewpoint independent reference frame with respect to the simulator world reference frame, and is therefore arbitrary.

Conclusions.

These results show that the rotation-compensated system is successful at estimating angular velocity in all the tests. The angular velocity error standard deviation compares well with the angular measurement noise and the position noise variance for the absolute state estimates is about 17 times smaller than the measurement noise. The small mean deviations show that the reference frame is accurate and, with the small variances, they show its temporal stability. The polar offset bias is the angular equivalent of the viewpoint independent reference frame origin -- it is arbitrary unless the viewpoint independent frame is required to align with an *a priori* map.

5.4.2 Linear Observer Motion.

The second series of tests comprised the three linear motion tests used in section 4.4.2 (page 106) for assessing the performance of the linear system (the set named LIN). The observer velocity was (0.08,0.06) metres per second, the square world was stationary, and

the observer angular velocity was zero. The world size in the three tests was respectively 40m, 70m and 100m. The noise parameters were identical with those used in section 4.4.2, i.e. 10 mrd standard deviation azimuth noise and 1cm standard deviation range noise. The sampling rate was 2.5Hz.

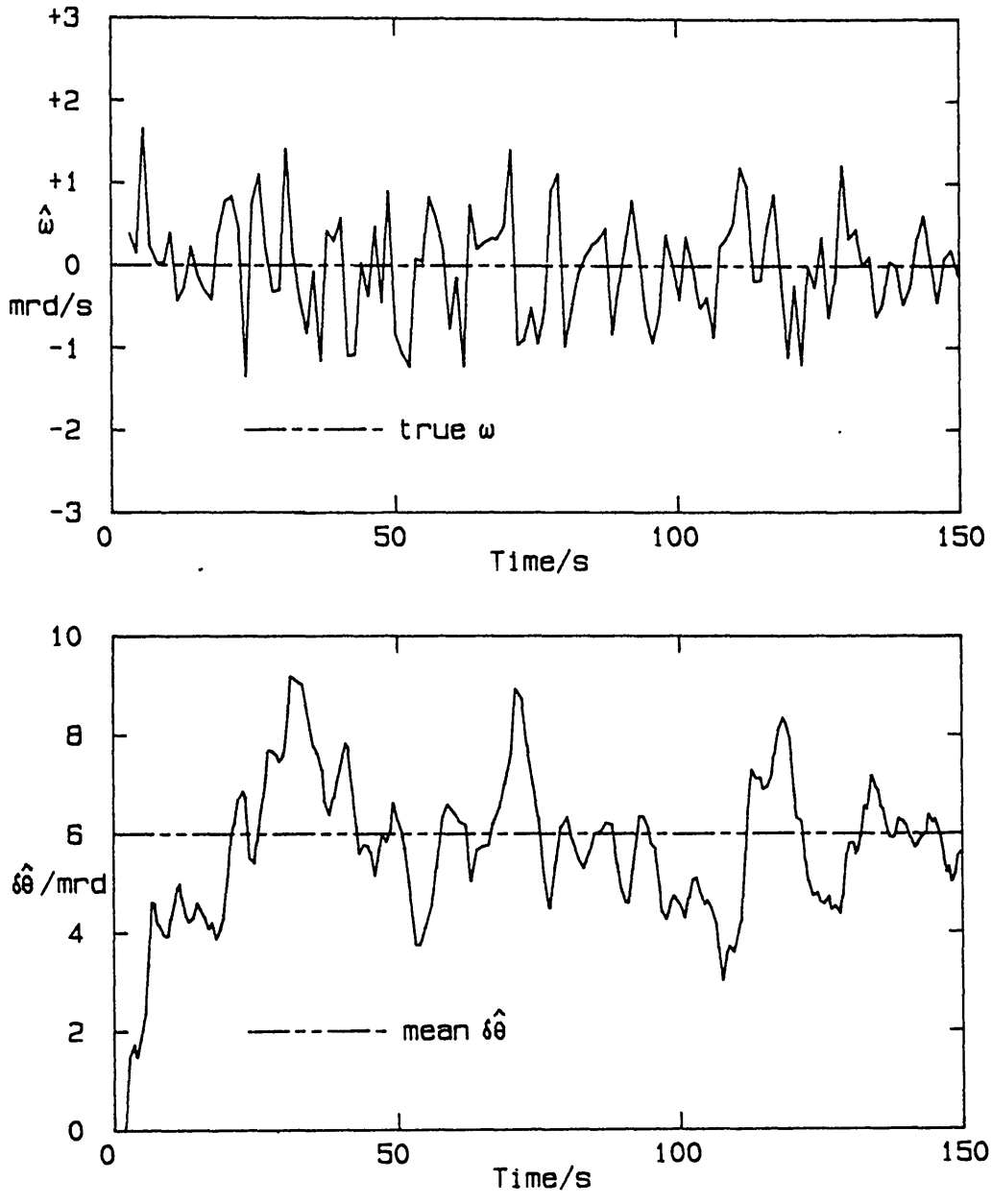
Results for these tests are presented in Tables 5.4 and 5.5 and in Figures 5.15 and 5.16. The figures illustrate the system performance in the 40m world test (lin1) using graphs of the observer angular velocity and polar offset error against time (Figure 5.15) and an x-y scatter plot of the observer and object absolute positions (Figure 5.16). The small scatter plots in the latter figure are magnifications of the scatter plot of the observer position and of the position of target 4.

Test	Mean $\hat{\omega}/\mu rds^{-1}$	Error Variance of $\hat{\omega}/mrd^2s^{-2}$	$\hat{\theta}/mrd^2$
lin1	4.72	0.400	1.489
lin2	5.91	0.175	1.197
lin3	6.91	0.105	1.008

Table 5.4. Angular Error Data for the LIN tests.

Table 5.4 gives the error analysis of the angular quantities for each test. In each test the true angular velocity and polar offset angle were zero. For all the tests the mean estimated angular velocity is small (a few microradians per second), and the small variances attest the good temporal stability of the viewpoint independent reference frame and the noise reduction properties of the system. The small mean angular velocity values testify to the velocity extraction performance.

In Table 5.5 the position and velocity error analysis results for the LIN tests are recorded and the affine stability indices are presented. The velocity error $\delta\underline{v}$ and position error variances (PEV) for the observer and each target are tabulated for each test. The data in this table complements the graphical presentation of Figure 5.16. In all the tests the observer velocity is extracted reliably



The actual value of the mean polar offset error is arbitrary (see text discussion on page 159).

Figure 5.15. Graphs of $\hat{\omega}$ and $\delta\hat{\theta}$ for test lin1

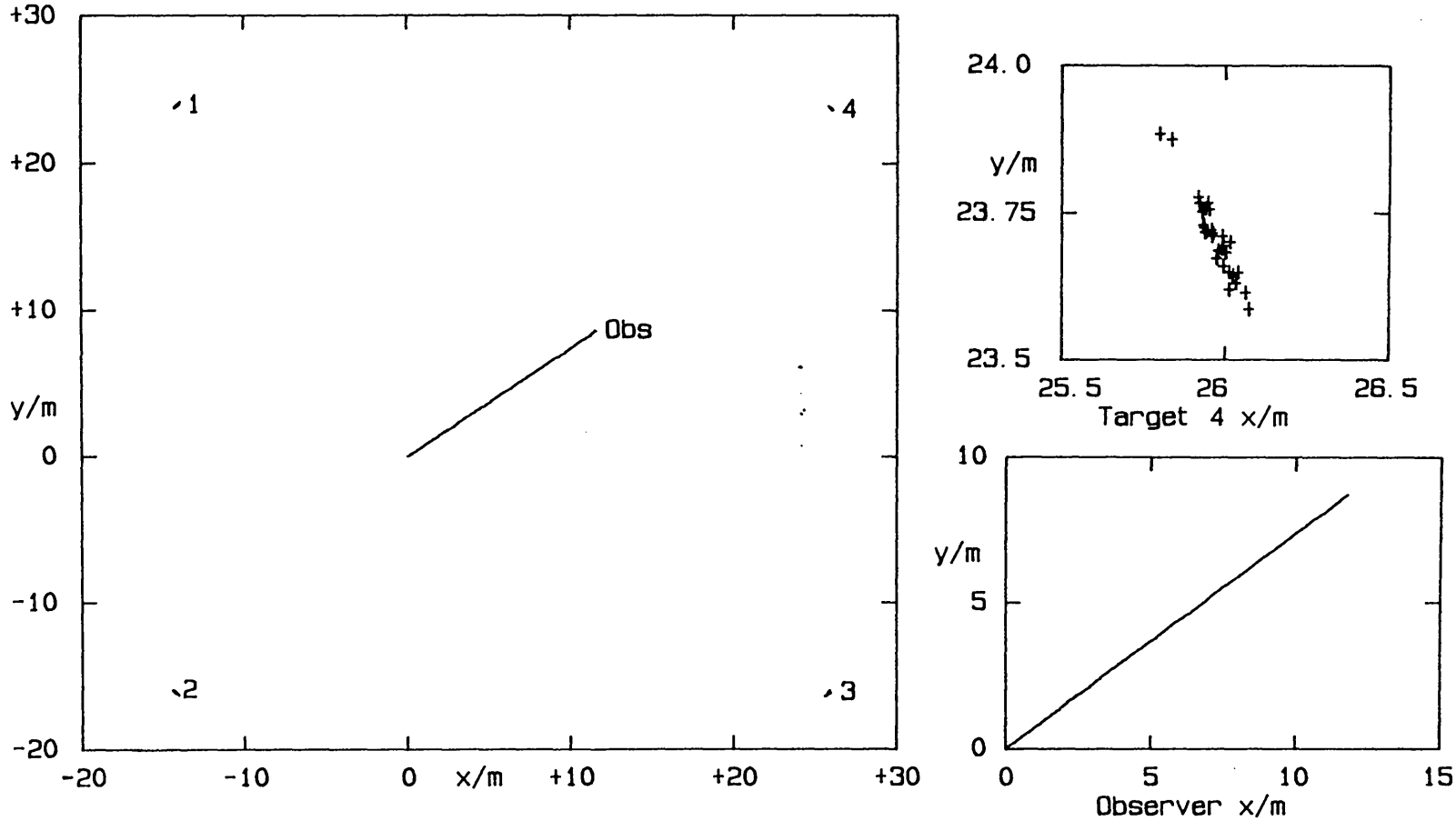


Figure 5.16. Scatter Plot of Absolute Target Positions in Test lin1.

Test	Target	Mean $\delta v/mms^{-1}$ x	y	PEV/cm^2 x	y
lin1	Obs	0.308	-0.348	3.190	6.038
	1	0.847	-0.301	83.72	85.47
	2	0.204	-0.256	78.21	92.23
	3	-0.931	-0.089	123.3	86.97
	4	-0.043	-0.091	49.68	54.61
lin2	Obs	0.315	-0.333	4.151	7.869
	1	1.033	-0.375	200.6	210.5
	2	-0.160	-0.412	190.5	207.2
	3	-0.952	-0.109	291.7	227.9
	4	0.407	-0.579	122.1	132.1
lin3	Obs	0.333	-0.456	4.794	9.338
	1	0.976	-0.284	345.6	366.3
	2	-0.486	-0.713	334.5	353.2
	3	-0.814	-0.023	510.7	419.8
	4	0.771	-0.983	218.2	232.6

Affine Transformation Analysis Results

Test	Mean $\delta\hat{\theta}$/mrd	Rotation/mrd	ε/%	λ_{min}/cm^2	λ_{max}/cm^2
lin1	6.00	5.79	0.01	0.01572	89.93
lin2	6.53	6.28	0.01	0.01348	90.75
lin3	6.88	6.62	0.01	0.01830	91.74

Table 5.5. Linear Error and Affine Analysis for the LIN tests.

to an accuracy of better than 0.6 mms^{-1} which, for the true observer speed of 10 cms^{-1}, is an error of better than 0.6%. The measurement noises in the tests had variances 800 cm^2 for lin1, 2500 cm^2 for lin2, and 9800 cm^2 for lin3, so the system shows a noise reduction factor of about 7 times in the first two tests and some 19 times in the last. The observer state noise variance is much smaller than the target variances, as might be expected since the observer estimates are generated by combining data from all four targets. The increase in the error variances with increasing target range is also clearly apparent in the table. This trend is a consequence of the predominantly azimuthal measurement errors.

The affine analysis results show the viewpoint independent reference frame to be stable in all the tests. Note that the units of the fit error eigenvalues are now square centimetres, so the affine fit

is somewhat worse than in the angular motion tests. The frame remains rigid (ε very small) and the correspondence between the affine rotation index and the mean polar offset error is good, indicating that the drift during the settling period is small.

Conclusions.

The performance of the system is also satisfactory in a situation where there is only linear observer motion. The observer position is tracked with better than 3.5 centimetre standard deviation, while the targets are tracked with a worst case standard deviation of about 23 cm (for test lin 3, target 3), in each position component. The stability of the reference frame, both temporal and spatial, is good, and the algorithm achieves a significant noise reduction.

5.4.3 The SR Monte Carlo Test Series.

The next series of tests was designed to exercise the observer linear and angular velocity extraction mechanisms simultaneously. All the targets in these tests were stationary, so no initialisation problems were anticipated, but the observer had non-zero linear and angular velocities.

The SR series comprised twenty Monte Carlo tests. The number of targets in each test was between two and ten, chosen at random with equal probability for each integer, and the target positions were distributed uniformly in a 200 metre square centred on the origin. The observer linear and angular velocity components were selected from zero mean Gaussian distributions with standard deviations of 10 cms^{-1} and 100 $mrds^{-1}$ respectively.

Each test ran for 120 second of simulated time, and the sonar noise parameters were set to represent a medium resolution device. As in section 4.2.2, the range resolution was ±3 cm and the angular resolution about 3.4°, i.e. 1 cm range and 10 mrd angular noise standard deviations.

Test	N	ω/mrds^{-1}	Mean $\hat{\omega}$/mrds^{-1}	Error Variance of $\hat{\omega}$/mrd^{2}s^{-2}	$\hat{\theta}$/mrd^2
sr1	4	181.0	181.7	0.155	2.707
sr2	4	-136.3	-136.1	0.145	5.352
sr3	3	-9.81	-9.76	0.169	3.074
sr4	10	108.5	108.5	0.759	1.847
sr5	7	60.85	60.89	0.427	1.661
sr6	9	-84.94	-84.91	0.246	1.438
sr7	9	44.99	45.01	0.060	0.713
sr8	10	62.13	62.13	0.529	2.194
sr9	6	-13.40	-13.37	0.273	3.110
sr10	9	-47.13	-47.11	0.285	1.409
sr11	4	15.29	15.36	0.150	2.808
sr12	3	88.50	88.60	0.902	4.717
sr13	7	19.83	19.86	0.567	2.648
sr14	3	-156.2	-156.2	0.913	3.457
sr15	7	2.65	2.69	0.088	1.160
sr16	4	27.26	27.27	2.045	5.231
sr17	5	18.12	18.15	0.321	2.075
sr18	7	79.44	79.47	0.398	1.967
sr19	4	80.38	80.37	4.192	9.175
sr20	4	-64.34	-64.27	0.142	2.946

Table 5.6. Angular Observer Motion Data for the SR Test Set.

Test	N	$\underline{v}$/mms^{-1} x	$\underline{v}$/mms^{-1} y	Mean $\delta\underline{v}$/mms^{-1} x	Mean $\delta\underline{v}$/mms^{-1} y	PEV/cm^2 x	PEV/cm^2 y
sr1	4	-94.06	40.60	1.324	0.181	13.73	11.72
sr2	4	57.76	287.3	-2.563	0.833	126.6	13.71
sr3	3	-113.0	-50.02	-0.406	0.560	3.17	4.93
sr4	10	-211.9	210.9	3.266	2.693	102.2	66.85
sr5	7	-61.28	-5.82	0.423	1.042	0.96	4.08
sr6	9	26.16	-157.0	0.481	0.317	3.12	0.55
sr7	9	-8.12	31.15	-0.226	-0.067	0.97	4.08
sr8	10	-51.81	128.2	1.070	0.467	21.42	5.09
sr9	6	65.76	132.5	0.635	-0.679	8.42	2.47
sr10	9	51.76	71.25	-0.607	0.380	2.87	0.65
sr11	4	145.2	32.77	0.295	-0.638	2.94	14.39
sr12	3	132.4	345.0	6.505	-1.225	352.3	23.35
sr13	7	73.45	-10.50	-0.279	-0.266	1.71	4.06
sr14	3	-57.53	91.59	-1.285	-0.634	21.67	7.73
sr15	7	-43.07	26.09	-0.597	0.322	3.89	0.34
sr16	4	-91.73	-55.36	-0.592	0.329	4.75	2.61
sr17	5	-2.86	-5.23	-0.279	-0.005	1.34	0.80
sr18	7	-37.36	54.62	0.294	0.255	5.78	4.17
sr19	4	151.9	-42.50	-0.207	-1.741	2.79	37.24
sr20	4	2.28	-8.53	-0.629	-0.843	5.79	10.40

Table 5.7. Linear Observer Motion Data for the SR Test Set.

Table 5.6 presents the observer angular motion data for the SR test set. The items tabulated are the number of targets, the true and mean estimated observer angular velocity, and the error variances of the observer angular velocity and polar offset angle estimates.

The results show a good angular velocity extraction performance. The maximum error is 0.7 $mrds^{-1}$ in test sr1, while the largest percentage errors are in tests sr15 and sr3 with 1.5% and 0.51% respectively. The percentage error is typically 0.15% to 0.3%.

The angular velocity variances, typically 0.2 to 0.5 mrd^2s^{-2}, show the long term accuracy and noise reduction properties of the system. The polar offset variances indicate that the observer orientation is estimated with a typical standard deviation of 1.2 to 1.8 mrd, which corresponds to a target position error of less than 2 mm per metre of range. The error variance of the angular velocity estimate tends to rise as the number of targets falls, because the estimate is a composite incorporating data from all the targets; however, there is considerable random variation in the computed variances from the tests.

In Table 5.7 the observer linear motion statistics for the SR test set are recorded. The table lists the true observer velocity $\underline{v}$, the mean observer velocity error $\delta\underline{v}$, and the error variances of the observer position estimates (the position error variances, PEV) for each coordinate dimension.

The relationship between the mean velocity error and the position error variance can be seen clearly in the table, for example in tests sr2, sr5, sr7 and sr12. Large velocity error tend to be associated with large position error variances, and the combination suggests temporal instability. The quality of the system linear velocity extraction performance is also apparent in the table, with most velocity errors being less than 1 mms^{-1} in size, and the observer position is estimated with a typical standard deviation of a few centimetres in each component.

The target error statistics for each test are tabulated in condensed form in Tables 5.8a and 5.8b. For each test the largest and smallest velocity error components were found and the value of those

Test	$\delta v/mms^{-1}$ Max. & Min.	Approx. Range/m	Measurement Variance/m^2	PEV/m^2 x	y
sr1	5.69 (max)	82	0.67	0.01595	0.04642
	0.01 (min)	58	0.34	0.03045	0.02409
sr2	4.80	61	0.37	0.05738	0.00165
	0.28	97	0.94	0.03593	0.1056
sr3	7.22	72	0.52	0.01096	0.04633
	0.09	58	0.34	0.00439	0.04852
sr4	13.16	96	0.92	0.01174	0.09500
	0.19	54	0.29	0.03110	0.00370
sr5	12.02	121	1.46	0.05791	0.08238
	0.20	17	0.03	0.00020	0.00491
sr6	6.28	68	0.46	0.00144	0.08819
	0.18	68	0.46	0.00144	0.08819
sr7	10.02	79	0.62	0.08948	0.00169
	0.05	38	0.14	0.00036	0.02584
sr8	6.68	83	0.69	0.00140	0.09280
	0.11	83	0.69	0.00140	0.09280
sr9	7.93	70	0.49	0.00295	0.1045
	0.39	106	1.12	0.08206	0.01699
sr10	11.75	101	1.02	0.00653	0.1283
	0.04	79	0.62	0.05155	0.02331

Table 5.8a. Target Motion Statistics for the SR Tests.

components, the range, measurement noise variance and position error variances for the appropriate targets are presented. The measurement noise is principally the azimuth noise, so the measurement noise variance is proportional to the square of the target range. Note that the unit of position error variance is ten thousand times larger in this table than in Table 5.7.

A comparison of the measurement and position error variance columns in Tables 5.8a and 5.8b shows the system noise reduction performance. The ratio of maximum PEV component to measurement noise is between 5 and 20. It is also apparent that large velocity errors tend to correlate with large PEV components. The individual target velocity errors may be fairly large (maximum 1.3 cms^{-1} in sr13), but the target position estimates have a typical standard deviation of 30 cm. Errors as small as 3.4 cm standard deviation (sr18) may be obtained. Thus the temporal stability of the reference frame appears

Test	δv/mms^{-1} Max. & Min.	Approx. Range/m	Measurement Variance/m^2	PEV/m^2 x	y
sr11	5.57 (max)	85	0.72	0.01323	0.1038
	0.06 (min)	91	0.83	0.05240	0.1001
sr12	2.26	71	0.50	0.02927	0.00243
	0.05	71	0.50	0.02927	0.00243
sr13	16.50	98	0.96	0.1740	0.01520
	0.03	13	0.02	0.00432	0.00153
sr14	0.45	123	1.51	0.04402	0.08509
	0.02	51	0.26	0.01739	0.00692
sr15	14.12	101	1.02	0.1310	0.02548
	0.14	85	0.72	0.06588	0.00125
sr16	1.87	92	0.85	0.00712	0.07435
	0.17	92	0.85	0.00712	0.07435
sr17	9.58	123	1.51	0.06979	0.04018
	0.14	24	0.06	0.00994	0.00268
sr18	9.47	100	1.00	0.07239	0.02312
	0.06	11	0.01	0.00129	0.00060
sr19	2.24	87	0.76	0.00150	0.1083
	0.40	31	0.10	0.01179	0.00416
sr20	6.91	107	1.14	0.06696	0.02898
	0.88	40	0.16	0.00148	0.01820

Table 5.8b. Target Motion Statistics for the SR Tests.

to be satisfactory.

The data in these tables is complemented by a graphical presentation in Figures 5.17 and 5.18 of the performance of the system for a typical test in the series (test sr10). Figure 5.17 comprises graphs of the observer angular velocity and polar offset estimates, with the true angular velocity marked on the former as a dashed line. The angular velocity graph shows well the relative tracking filter initialisation period and the subsequent angular velocity transient as the absolute state estimation cycle begins. The angular velocity has settled down by t = 20 seconds and fluctuates at random about the true value. The polar offset also stabilises, fluctuating about the arbitrary drift value. Dashed lines on this graph indicate the ±1 standard deviation interval with respect to the mean polar offset error.

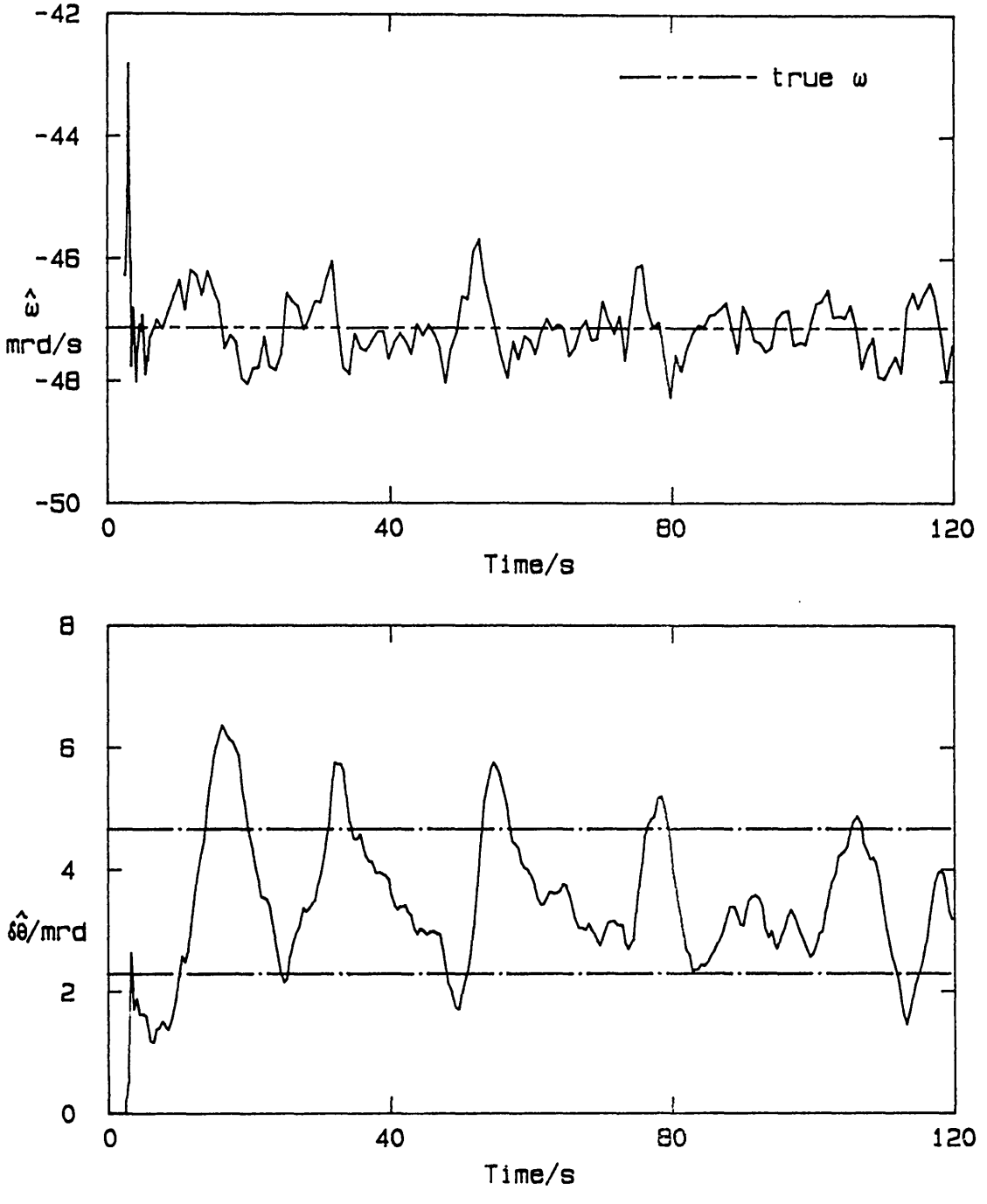


Figure 5.17. Graphs of $\hat{\omega}$ and $\delta\hat{\theta}$ for test sr10.

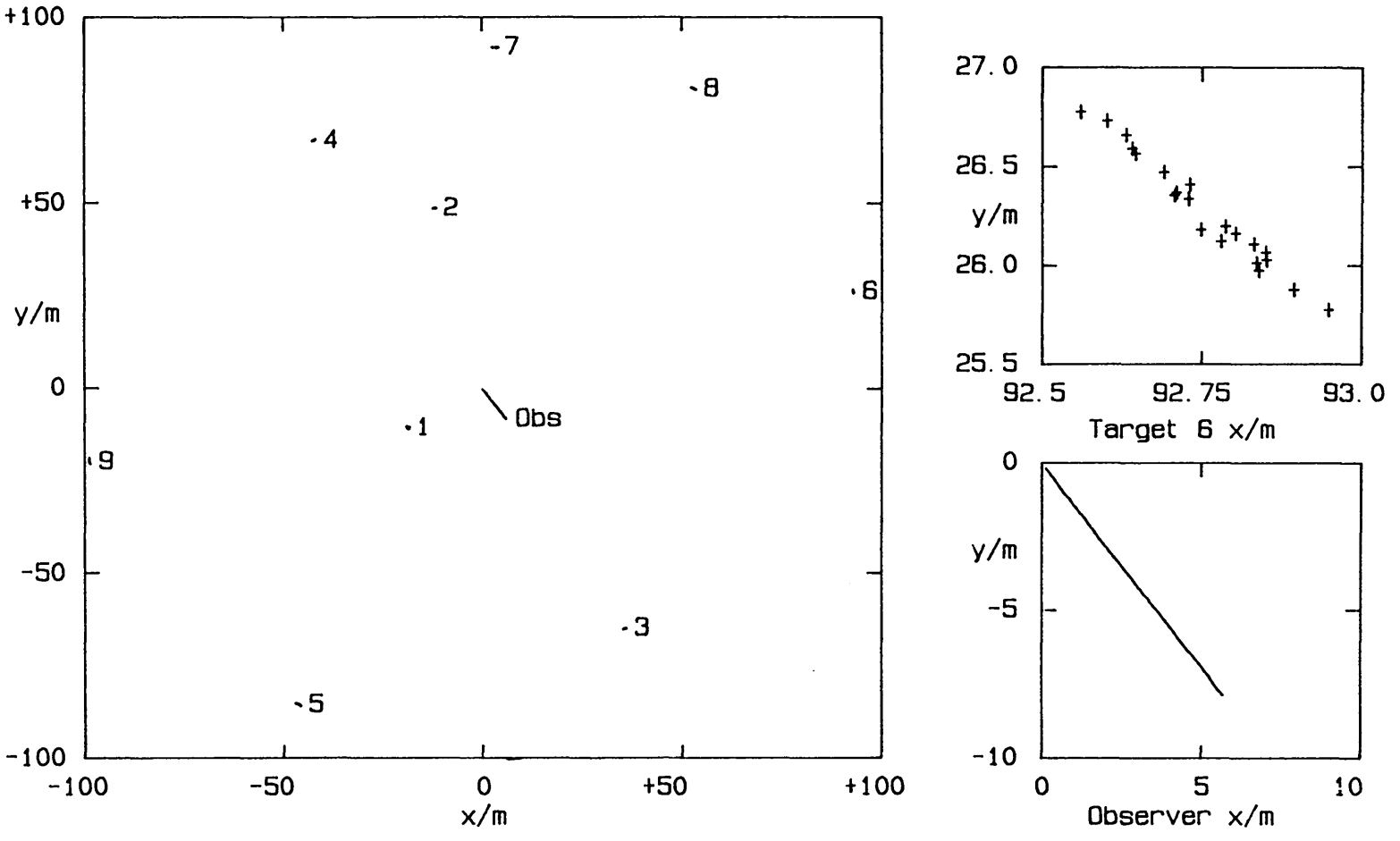


Figure 5.18. Scatter Plot of Target Absolute Positions for Test sr10.

Figure 5.18 is a scatter plot of the estimated target and observer absolute positions. The small inset figures are magnifications of the observer trajectory and of the scatter plot for target 4. The overall observer trajectory is correct and the targets exhibit no large scale movement, points being deposited randomly within the scatter region once the settling period is over.

Test	Mean $\delta\hat{\theta}$/mrd	Rotation/mrd	ε/%	λ_{min}/cm^2	λ_{max}/cm^2
sr1	6.83	7.77	-0.04	0.187	8.997
sr2	9.57	-10.36	-0.19	40.04	548.5
sr3	7.09	6.00	-0.01	0	7.284
sr4	0.88	16.04	0.07	233.1	786.1
sr5	5.03	12.62	-0.01	40.56	155.6
sr6	3.16	-6.46	0.07	218.8	329.8
sr7	3.49	8.78	0.01	16.56	34.03
sr8	1.78	10.96	-0.01	54.33	98.20
sr9	7.35	5.64	0.03	2.74	9.39
sr10	3.48	-1.82	-0.03	12.66	44.29
sr11	7.48	8.77	-0.01	1.829	19.69
sr22	9.19	15.84	0.27	0	229.1
sr13	2.85	5.29	0.01	2.00	15.70
sr14	2.52	-13.64	2.86	0	83.32
sr15	4.21	4.43	-0.01	1.694	2.756
sr16	2.30	5.36	-0.01	5.922	11.55
sr17	3.97	6.37	-0.01	7.247	24.80
sr18	3.31	11.36	0.13	25.78	69.39
sr19	8.61	14.66	0.24	0.05	143.1
sr20	7.38	-0.60	-0.14	13.00	3.09

Table 5.9. Affine Analysis Indices for the SR Tests.

The graphical presentation shows clearly the stability of the reference frame and the quality of velocity extraction performance achieved by the system. This judgement is vindicated by the results of the affine transformation analysis, which are presented in Table 5.9. The difference between the mean polar offset error and the average rotation of the affine transformation varies widely in these data, showing that the polar offset origin does drift by a significant amount during initialisation if the observer is rotating. However, it does not drift very much after initialisation since the polar offset error variances (in Table 5.6) are small.

Conclusions for the SR Tests.

The system performance in these tests was good. The algorithm stabilised the viewpoint independent frame, typically to within a few centimetres in each dimension. The frame orientation was also stable. The velocity extraction processing successfully separated the observer linear and rotary motion with as few as three targets and the velocity errors were small. Noise reduction and frame stability were satisfactory and as expected there were no initialisation problems, the system having settled down within twenty seconds of the start of the test.

5.4.4 The MN Monte Carlo Test Series.

The second series of Monte Carlo tests, named MN, used situations which included both stationary and moving targets. Each of the twenty tests in this series included four stationary reference targets (the system was not told that they were stationary, however) arranged in a 70m square centred on the origin, and also between one and six moving targets. These moving targets were uniformly distributed in a 200m square centred on the origin and their velocity components were generated using zero mean Gaussian random numbers with 20 cms^{-1} standard deviation. The number of targets was a uniform random variable and the observer linear velocity components were Gaussian with zero mean and 10 cms^{-1} standard deviation. The observer angular velocity was zero in all these tests, and each test ran for 120 seconds of simulated time. The same sonar parameters were used as in the SR series.

Tables 5.10 and 5.11 record the angular and linear statistics for the tracked observer motion in each test. Table 5.10 tabulates the mean angular velocity error (the true angular velocity is zero), the angular velocity and polar angle offset error variances, and the number of targets present in the test written as moving/total.

Test	N	Mean $\hat{\omega}$/mrds^{-1}	Error Variance of $\hat{\omega}$/mrd^{2}s^{-2}	$\hat{\theta}$/mrd^2
mn1*	3/7	-0.70	0.926	435.2
mn2	3/7	-0.28	2.210	104.5
mn3*	4/8	-0.93	1.844	969.3
mn4	2/6	-0.28	1.293	162.6
mn5*	5/9	1.09	2.432	1339.7
mn6	3/7	-0.04	1.935	5.1
mn7	4/8	0.05	1.193	12.3
mn8*	3/7	-1.44	0.769	2441.1
mn9	1/5	0.08	1.126	5.5
mn10	3/7	0.06	1.300	8.9
mn11	4/8	0.23	0.832	113.4
mn12	2/6	0.09	0.401	14.3
mn13*	3/7	1.16	0.909	1615.2
mn14	5/9	0.42	0.357	218.6
mn15	4/8	0.00	1.592	7.0
mn16	3/7	0.176	1.060	16.9
mn17*	1/5	2.218	0.891	4810.3
mn18	3/7	-0.116	0.877	39.7
mn19	3/7	0.527	0.232	291.6
mn20	1/5	0.171	0.497	11.0

Table 5.10. Angular Observer Motion Data for the MN Test Set.

Of the twenty tests, six can be considered failures on the basis of the angular velocity and polar offset estimation alone. They are marked with an asterisk in the tables. Nine tests gave good angular velocity extraction performance, with small angular velocity errors and low polar offset variances. The remainder had polar offset errors of more than 10 mrd standard deviation.

Turning to the linear motion data in Table 5.11, a further three tests exhibit severe velocity extraction errors. However, the remaining eleven tests achieve position error standard deviations of less than 20 cm and velocity errors of typically 5 mms^{-1}.

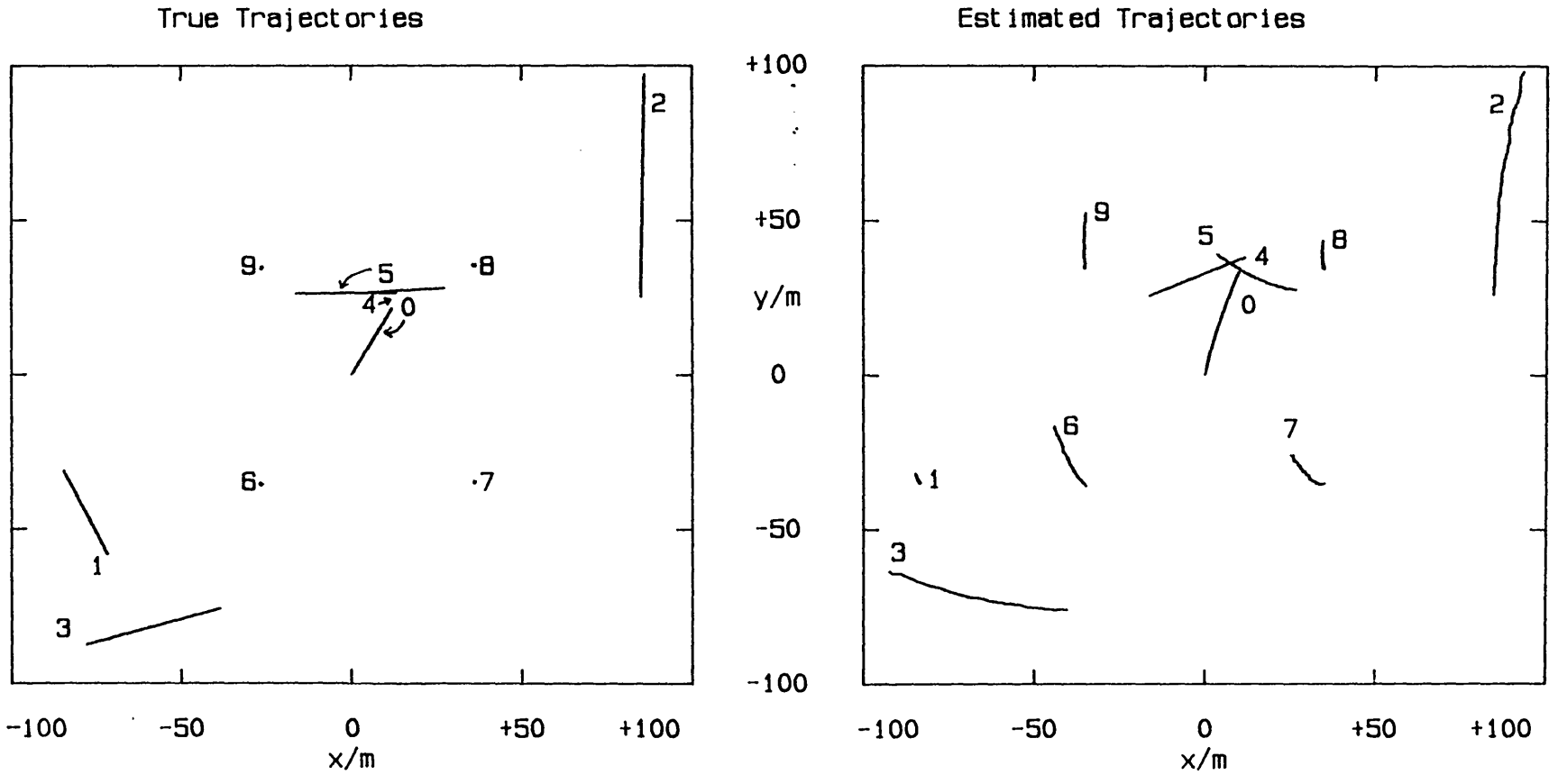
The reason for the failure of the asterisked tests is made clear in graphical presentation. Figure 5.19 displays in scatter plot form the true and estimated absolute positions for the targets and observer in test mn5. Each scatter track is labelled with the target number; the observer is target zero. The problem in the test is that a target (number one) has been misclassified as stationary. The system has therefore stabilised its position at the cost of biassing the

		$\underline{v}$/mms^{-1}		Mean $\delta\underline{v}$/mms^{-1}		PEV/m^2	
Test	N	x	y	x	y	x	y
mn1*	3/7	33.81	-171.7	119.8	-15.21	16.52	0.29
mn3*	4/8	-64.21	-177.5	13.60	-19.63	0.34	0.49
mn5*	5/9	100.7	181.0	12.25	-107.4	0.29	12.02
mn8*	3/7	-77.78	88.19	-42.71	66.94	1.73	15.12
mn13*	3/7	211.0	-64.29	31.07	79.50	0.99	6.33
mn17*	1/5	-11.96	161.7	38.35	65.99	1.54	4.13
mn2	3/7	-56.93	21.08	28.90	-33.08	1.02	1.36
mn11	4/8	-9.34	45.65	39.68	40.21	1.38	1.53
mn19	3/7	40.51	115.9	58.21	10.90	3.90	0.13
						PEV/cm^2	
mn4	2/6	-95.14	53.95	1.24	0.07	35.88	18.45
mn6	3/7	208.5	225.3	-2.40	-0.53	38.00	13.87
mn7	4/8	64.45	18.09	3.84	-4.23	125.2	107.9
mn9	1/5	-52.97	84.54	1.22	0.38	14.50	2.88
mn10	3/7	140.7	-84.30	3.18	-0.53	101.1	7.86
mn12	2/6	153.8	89.30	-5.00	-5.76	274.6	400.1
mn14	5/9	10.08	36.29	-40.27	2.31	140.7	42.37
mn15	4/8	53.99	-131.8	3.41	4.26	36.60	188.4
mn16	3/7	-57.53	91.59	-2.28	-0.38	68.09	24.27
mn18	3/7	-11.97	-202.8	-5.41	1.36	282.1	35.62
mn20	1/5	-36.34	13.03	0.73	0.57	11.53	74.51

Table 5.11. Linear Observer Motion Data for the MN Test Set.

velocity estimates of all the other targets in the test. All the tests that failed suffer from this problem, either remaining in their start-up initialisation phase or re-entering an initialisation phase at some point during the test run.

Table 5.12 presents the velocity extraction and noise reduction statistics for those tests deemed to have succeeded on the basis of their observer motion tracking performance. The left hand section of the table gives the statistics for the moving targets in each test, the right gives those for the four stationary reference targets. As before, the maximum and minimum velocity error components were found and the error statistics tabulated for those targets. For moving targets the velocity error component, the size of the target velocity, the approximate target range, and the appropriate (maximum or minimum) position error variance are given; for stationary targets the velocity error component and position error variance are stated (these targets are always about 50 m from the observer).



Targets move towards their numbers; the observer is target 0.

Figure 5.19. Scatter Plot of Absolute Target Positions for Test mn5.

Test	N	Moving $\delta\underline{v}$/mms^{-1}	Moving $\lvert\underline{v}\rvert$/mms^{-1}	Moving Range/m	Moving PEV/m^2	Stationary $\delta\underline{v}$/mms^{-1}	Stationary PEV/m^2
mn4	2/6	23.08 (max) 14.85 (min)	93.03	74	0.697 0.351	20.0 0.37	0.182 0.173
mn6	3/7	10.83 0.37	112.4 278.3	97 106	0.214 0.040	5.49 0.82	0.050 0.047
mn7	4/8	12.56 0.03	222.9 218.3	65 88	0.181 0.016	6.25 0.13	0.060 0.019
mn9	1/5	1.34 0.83	58.33	59	0.092 0.001	5.69 0.16	0.039 0.016
mn10	3/7	4.78 0.22	132.9 247.4	96 74	0.162 0.054	9.77 1.51	0.074 0.021
mn12	2/6	7.51 0.29	136.9 194.9	46 23	0.038 0.018	5.27 0.33	0.113 0.019
mn14	5/9	94.0 2.82	73.9 333.5	111 21	1.467 0.156	6.46 0.62	0.301 0.223
mn15	4/8	9.68 1.26	19.32 479.2	96 58	0.140 0.123	10.92 1.49	0.056 0.018
mn16	3/7	23.8 2.39	131.6 326.1	93 82	0.320 0.048	5.19 0.27	0.048 0.035
mn18	3/7	28.14 6.92	38.71 61.92	88 65	0.314 0.026	12.71 3.93	0.089 0.052
mn20	1/5	2.35 1.78	174.3	61	0.059 0.005	8.45 1.22	0.073 0.025

Table 5.12. Target Motion Statistics for the MN Test Set.

There are three things to note in this table. First, the minimal velocity errors tend to be associated with the high speed targets. This is to be expected since such targets are most easily identified as moving. Second, it is possible for a reference target to have a large velocity error, with correspondingly large variances, without seriously upsetting the stability of the reference frame. This is exemplified in test mn15. The large error associated with the stationary target suggests that it was misclassified as moving (and therefore able to drift freely). Third, the system does achieve a noise reduction for all the stationary targets except in test mn14 (the measurement noise variance for this target is about 0.25 m^2) and for moving targets in all tests except mn4 and mn14.

Test	Mean $\delta\hat{\theta}$/mrd	Rotation/mrd	ε/%	λ_{min}/cm^2	λ_{max}/cm^2
mn4	2.21	4.89	0.04	1.87	17.56
mn6	4.02	4.01	-0.04	11.31	32.97
mn7	7.31	7.17	0.02	5.23	22.83
mn9	9.64	9.50	0.00	4.85	10.41
mn10	9.67	9.39	0.06	18.19	42.46
mn12	16.53	16.00	-0.01	28.18	120.7
mn14	26.33	25.79	0.11	94.39	422.1
mn15	-3.43	-4.11	0.00	0.36	2.72
mn16	17.36	15.05	0.12	54.63	290.3
mn18	-23.18	-22.92	-0.06	71.02	142.4
mn20	11.64	11.30	0.07	1.3	20.50

Table 5.13. Affine Analysis Indices for the MN Tests.

These data indicate that the reference frame is stable over time in these tests, but the system performance is not as good as in the simpler cases reported in sections 5.4.2 and 5.4.3. Affine transformation analysis data, presented in Table 5.13, shows that the reference frame is spatially stable and that there is relatively little polar offset drift during the initialisation period.

Finally, Figures 5.20 and 5.21 present graphically the results of test mn15. The former figure graphs the estimated polar offset angle observer angular velocity against time, while the latter shows the true and estimated absolute positions for the observer and targets as x-y scatter plots. Figure 5.22 shows magnifications of the observer track and of the tracks of target numbers three, four and five. For the moving targets the tracks are correct and develop as indicated by the arrows; the points for the stationary target the points are deposited randomly within the scatter region.

Conclusions for the MN Tests.

The major conclusion to be drawn from this test set is that the motion resolution system works well provided that it succeeds in establishing a correct initialisation. Under adverse conditions (many moving targets) the system succeeded in eleven out of twenty cases; its failures were generally due to a target being misclassified as stationary at some point during the test. The performance is

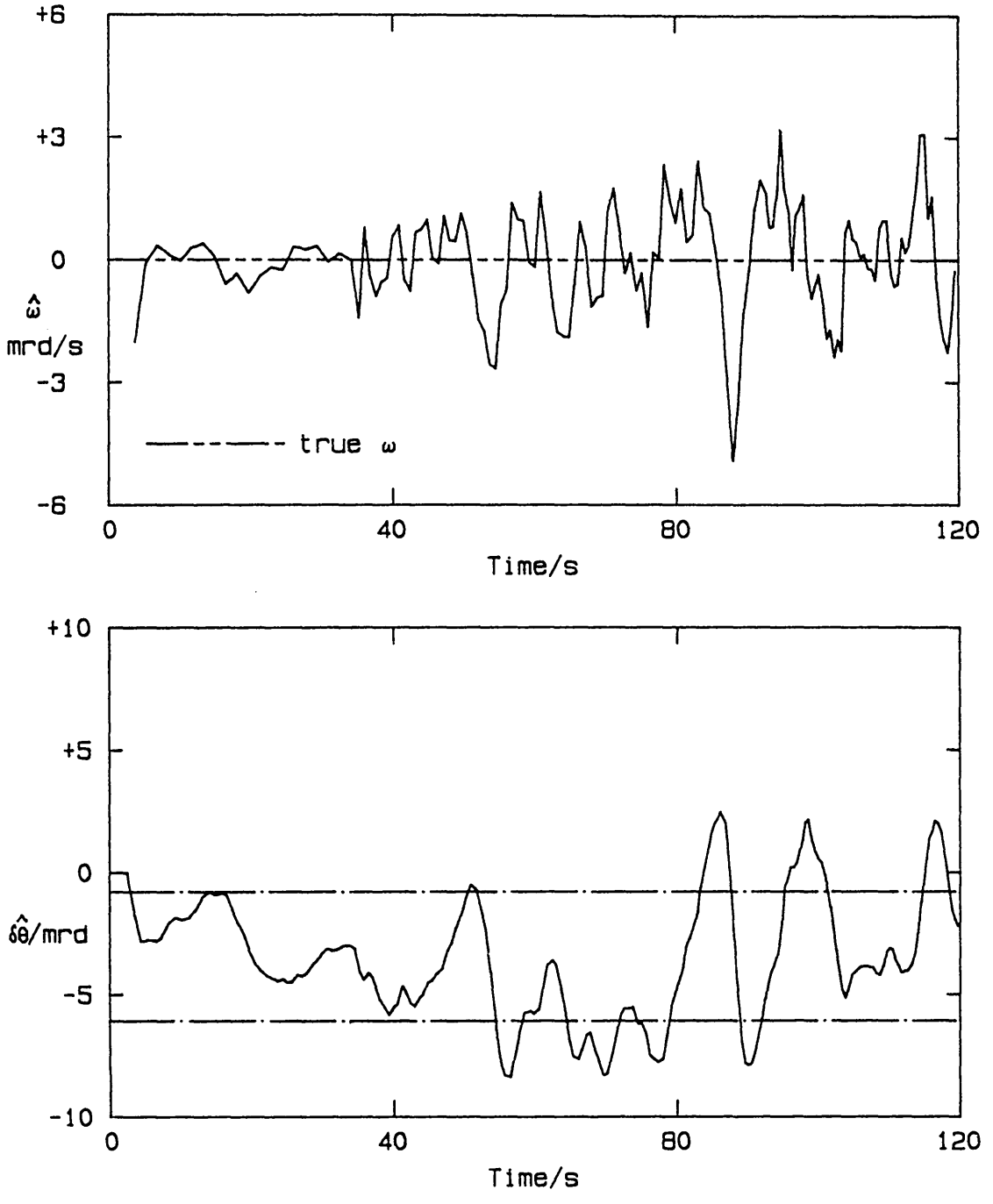


Figure 5.20. Graphs of $\hat{\omega}$ and $\delta\hat{\theta}$ for test mn15.

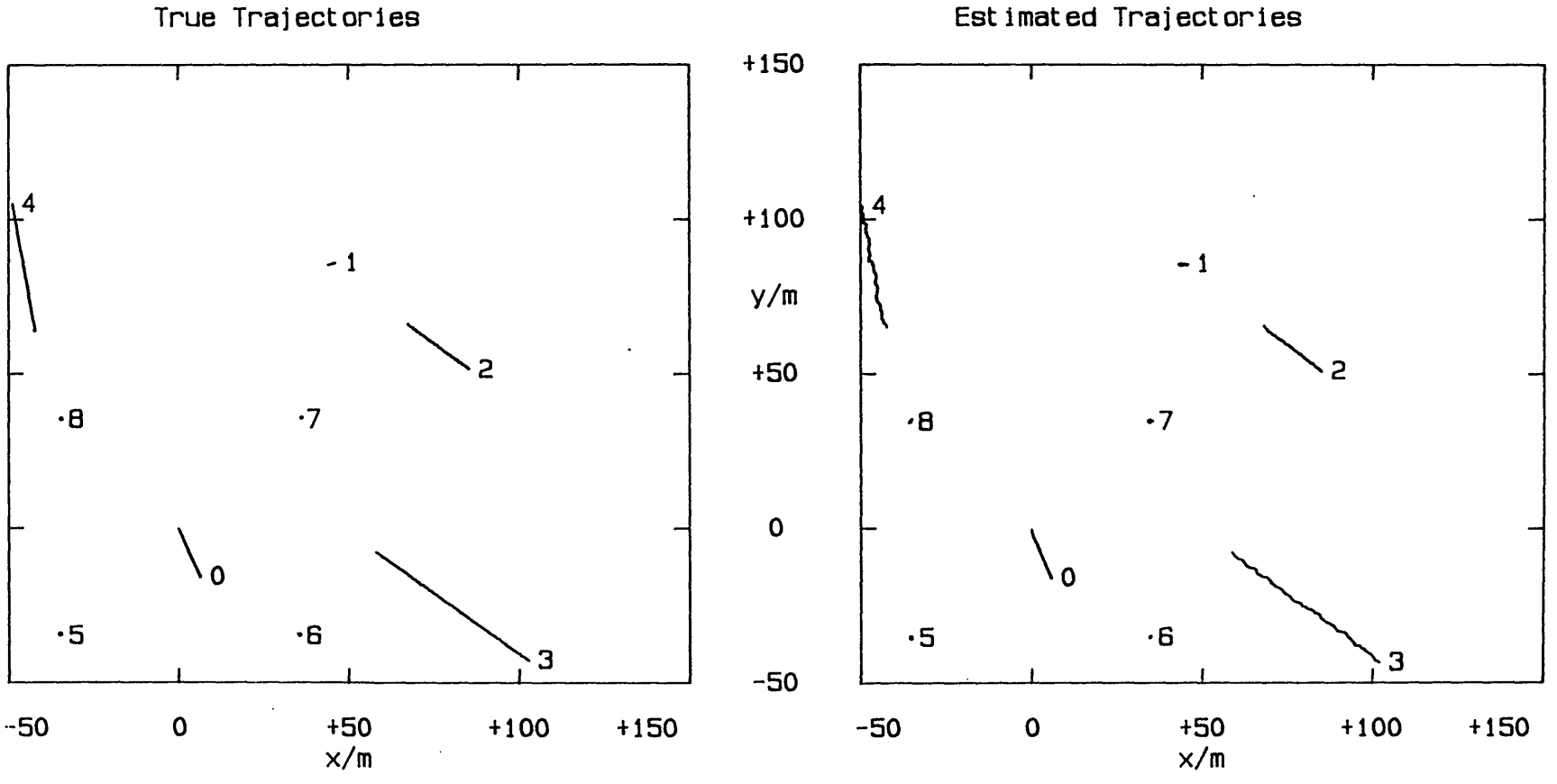


Figure 5.21. Scatter Plot of Absolute Target Positions for Test mn15.

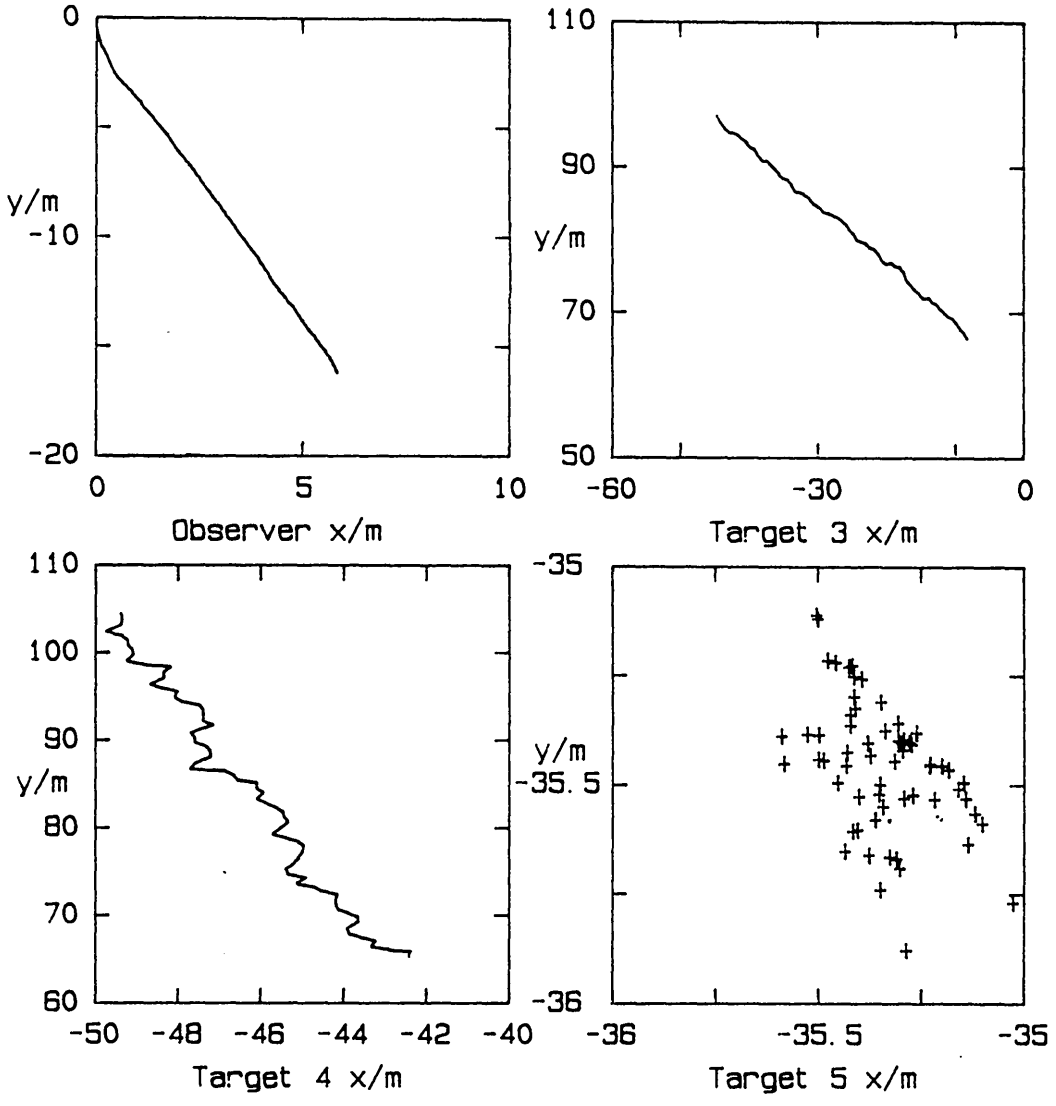


Figure 5.22. Magnified Regions of Figure 5.21.

not as good as in the simpler cases, but this is reasonable in view of the greater likelihood of stationary targets being misclassified (because the moving targets increase the velocity errors in the system) and the initialisation phase at the start of these tests (the earlier tests did not have an initialisation phase since there were no moving targets to misclassify).

5.4.5 The MR Monte Carlo Test Series.

The last set of Monte Carlo tests was identical in design to the MN series except that the observer angular velocity was a Gaussian random variate with zero mean and 100 mrds^{-1} standard deviation.

Twenty tests were run in this series of moving target and rotating observer tests.

Test	N	ω/mrds^{-1}	Mean $\hat{\omega}$/mrds^{-1}	Error Variance $\hat{\omega}$/mrd^{2}s^{-2}	$\hat{\theta}$/mrd^2
mr1	4/8	-68.34	-67.78	0.567	384.4
mr2	5/9	22.23	22.01	1.435	79.1
mr3*	4/8	-8.23	-10.86	11.12	5649.0
mr4	3/7	94.04	93.85	1.578	21.5
mr5*	4/8	105.5	102.7	1.967	8423.0
mr6	1/5	47.40	48.06	1.249	329.3
mr7	5/9	79.22	79.19	0.571	2.8
mr8	1/5	-41.49	-41.18	0.924	67.8
mr9*	4/8	29.42	30.56	1.810	1213.7
mr10*	5/9	-11.18	-12.61	25.10	4403.1
mr11*	3/7	117.0	115.8	5.042	1640.0
mr12	1/5	-100.5	-100.2	1.035	40.0
mr13*	4/8	62.67	63.90	1.285	1506.0
mr14	1/5	-100.0	-99.83	0.789	15.5
mr15*	5/9	-48.62	-46.17	1.596	6625.3
mr16*	2/6	-3.46	-1.89	1.516	2213.5
mr17	5/9	-45.60	-45.08	1.404	353.5
mr18*	1/5	-37.35	-38.67	3.437	1897.6
mr19	1/5	-187.1	-187.1	1.136	7.0
mr20	2/6	8.48	8.63	1.314	26.8

Table 5.14. Angular Observer Motion Data for the MR Test Set.

The angular and linear observer motion data are presented in Tables 5.14 and 5.15 respectively. The major difference between these tests and the MN series is that the proportion of failures is higher, reflecting the greater complexity of the situations used for testing. Nine tests were failed on their angular estimation performance; it appears that any test with a mean angular estimation error exceeding about 0.5 mrds^{-1} is likely to fail. The angular velocity estimation error variances were typically slightly higher than in the MN tests. The linear motion data in Table 5.15 provides evidence to fail another three tests which suffer from substantial observer velocity errors. The eight tests that remain, however, show fairly good velocity extraction performance and achieve observer errors of 25 cm, or less, standard deviation (except for mr6 and mr2).

It is much more obvious in these tables that the tests with a high proportion of moving targets tend to fail. As before, the reason is

Test	N	$\underline{v}$/mms^{-1} x	$\underline{v}$/mms^{-1} y	Mean $\delta\underline{v}$/mms^{-1} x	Mean $\delta\underline{v}$/mms^{-1} y	PEV/m^2 x	PEV/m^2 y
mr3*	4/8	-165.9	-61.28	-5.82	-10.56	0.014	0.133
mr5*	4/8	-84.94	-176.5	-130.4	-75.50	15.81	5.29
mr9*	4/8	88.25	-169.3	41.16	-3.29	2.75	0.036
mr10*	5/9	-12.52	52.74	-53.74	-43.01	4.16	3.06
mr11*	3/7	-151.9	-54.26	12.04	65.47	0.175	4.37
mr13*	4/8	63.24	-35.78	-15.78	-16.56	0.219	7.84
mr15*	5/9	102.8	-38.29	110.9	67.74	11.84	4.75
mr16*	2/6	-171.9	-33.06	88.29	29.66	7.11	0.922
mr18*	1/5	-56.93	21.08	-72.02	-46.69	5.33	2.27
mr1	4/8	114.4	-47.93	11.41	14.80	0.140	0.189
mr8	1/5	46.54	134.2	-50.00	10.99	2.62	0.093
mr17	5/9	87.68	110.4	-65.70	-97.83	4.42	9.98
						PEV/cm^2	
mr2	5/9	108.5	-157.0	10.21	-0.327	2275	15.3
mr4	3/7	-37.06	-22.04	-1.75	-0.26	146.4	86.3
mr6	1/5	-63.25	83.43	-7.11	1.85	705.4	75.4
mr7	5/9	-4.28	-26.00	0.27	1.31	6.3	8.5
mr12	1/5	108.8	106.7	1.49	2.45	80.9	213.6
mr14	1/5	46.71	60.69	-4.02	2.32	119.1	50.0
mr19	1/5	-63.19	-40.79	-2.00	1.19	38.9	21.4
mr20	2/6	39.95	-55.25	-2.85	-3.29	163.3	182.1

Table 5.15. Linear Observer Motion Data for the MR Test Set.

that moving targets are misclassified in failure tests and there is a higher probability of initialisation failure for a higher proportion of moving targets.

Tables 5.16 and 5.17 contain the target velocity extraction and noise reduction data and the affine transformation analysis results for those tests judged to have succeeded on the basis of the observer motion analyses. In Table 5.16 it is not so apparent as for the MN tests that the small velocity extraction errors occur with large velocities. The majority of the tests exhibit noise reduction, in some cases by a large factor, and the reference frames generated in the tests are metrically consistent.

Test	N	Moving $\delta\underline{v}/mms^{-1}$	Moving $\lvert\underline{v}\rvert/mms^{-1}$	Moving Range/m	Moving PEV/m^2	Stationary $\delta\underline{v}/mms^{-1}$	Stationary PEV/m^2
mr2	5/9	23.54 (max)	597.9	94	0.223	9.54	0.079
		0.94 (min)	257.5	80	0.032	0.64	0.079
mr4	3/7	14.77	120.4	101	0.202	11.01	0.057
		5.91	297.6	44	0.020	0.34	0.039
mr6	1/5	1.47	146.3	62	1.014	28.15	0.554
		1.08			0.004	0.33	0.188
mr7	5/9	9.64	75.32	101	0.116	5.48	0.018
		0.19	66.00	74	0.020	0.32	0.018
mr12	1/5	3.55	28.79	97	0.545	11.48	0.053
		1.58			0.038	0.09	0.034
mr14	1/5	5.19	293.0	36	0.056	10.75	0.039
		4.76			0.018	0.41	0.039
mr19	1/5	4.24	109.7	38	0.014	8.30	0.040
		2.05			0.010	0.15	0.030
mr20	2/6	34.06	549.5	88	0.512	19.51	0.177
		1.23	389.7	96	0.099	2.11	0.018

Table 5.16. Target Motion Statistics for the MR Tests.

Test	Mean $\delta\hat{\theta}$/mrd	Rotation/mrd	ε/%	λ_{min}/cm^2	λ_{max}/cm^2
mr2	-0.084	3.62	-0.02	16.17	39.41
mr4	-18.33	-8.18	-0.03	181.3	235.9
mr6	31.43	30.36	-0.23	57.96	155.8
mr7	-3.58	4.03	0.01	67.72	97.83
mr12	15.34	5.50	0.01	14.76	398.2
mr14	19.27	13.22	0.12	0.6103	78.30
mr19	8.31	-2.33	0.02	8.036	11.73
mr20	24.49	26.31	0.97	175.9	681.9

Table 5.17. Affine Analysis for the MR Tests.

Conclusions for the MR Tests.

This series of tests confirmed the conclusion that the principal cause of failure in the tests is initialisation failure, with re-entry into an initialisation phase as a secondary cause. If a moving target is misclassified during the operation of the algorithm, it tends to be held stationary, thereby forcing all the other targets to move. This situation is aggravated by the small number of stationary targets in the tests, as the information contributed by each is

significant in the computation.

However, despite the initialisation problem, the system was successful in eight of the twenty test, maintaining a consistent, stable reference frame and achieving a moderate noise reduction over the raw measurement noise.

5.5 Conclusions and Suggestions for Further Research.

This section summarises the conclusions and results from the tests reported in section 5.4, and concludes the discussion of two-dimensional motion resolution in this dissertation by giving a number of suggestions for future research.

5.5.1 A Summary of the Conclusions of Section 5.4.

The major conclusions to be drawn from the tests described above are as follows.

1. The motion resolution system is able to estimate observer rotation and linear motion accurately, either individually or together, in situations where all the targets seen are stationary. In those situations the system achieves a reasonable noise reduction, maintains a stable and accurate viewpoint independent reference frame, and suffers no initialisation problems. The initialisation of the system requires about 50 sightings of four targets.

2. The performance of the system in the presence of moving targets is also good, provided that the system initialises the viewpoint independent frame successfully. In that case, the velocity extraction, noise reduction, and frame stability properties are comparable with the situation where there are only stationary targets.

3. However, with linear observer motion and moving targets the system initialised successfully in only 55% of the tests, while when the observer was also rotating the initialisation succeeded in 40% of the tests. The initialisation failures, which were expected from the performance of the linear system in the ML tests of Chapter four, were in this case due entirely to the misclassification of moving targets.

The overall conclusion is that the major deficiency in the full two-dimensional motion resolution system is in the hypothesis testing mechanism responsible for identifying moving targets. As in Chapter four, a very simple hypothesis test was used. The mean absolute velocity and the absolute velocity covariance matrix were estimated from the sequence of target absolute velocity estimates generated by the system. If the mean velocity was significant with respect to its covariance then the 'moving' target hypothesis was accepted. An conservative ad hoc threshold of 1.2 (the expected value of the χ^2 statistic is 2) was found by experiment to be reasonable.

The motion resolving system is tolerant of misclassified stationary targets since these cannot cause an initialisation phase. However, it is intolerant of misclassified moving targets, since it attempts to make these stationary in the viewpoint independent reference frame and by doing so causes the frame to drift. The success of the system in those cases where the initialisation succeeds implies that the system performance could be greatly improved by the use of a more sophisticated hypothesis test (some suggestions are given in the next section).

There are two further points to make concerning the system initialisation performance. First, the systems described above are attempting to deduce the motion of the observer using only the information present in their sensory input. In practice, direct estimates of observer motion are generally available from the control system of the robot vehicle and the use of these would facilitate the system initialisation by providing an approximate initial reference frame. Second, the system performance in the tests described is hampered by the small number of targets simulated. With only four stationary

targets the contribution of each is important in the maintenance of the reference frame; a misclassification of one of these targets results in the loss of about a quarter of the information available to the system. In a practical application the number of targets available to the system would be much greater and the importance of each individual stationary target correspondingly less.

5.5.2 Some Suggestions for Future Research.

There are four major areas where the motion resolving system described above may be improved, and there is considerable scope for further testing, especially in the area of noise reduction for perturbed observer and object motion.

Extension to Three Dimensions.

The most obvious extension of the two-dimensional system is to enable it to handle full three-dimensional motion of the observer. This extension is described in detail in Chapter six.

Using Direct Observer Motion Information.

The motion resolution systems described have attempted to deduce the motion of the observer using only those cues present in sonar input. In most small robotic systems, however, direct estimates of the vehicle motion are available from the control and navigation systems of the robot. These could be used to improve the initialisation performance of the motion resolution system. Techniques for incorporating direct observer motion estimates are discussed in Chapter six.

Improved Hypothesis Testing.

There are numerous possibilities for improved hypothesis testing to determine the motion status of targets. Some examples are given here.

- More sophisticated statistical tests, such as distribution tests on the estimated velocities, may be tested.

- The behaviour of targets over time may be exploited. Targets that have been found to be stationary for an extended period may be assumed to be fixed in the environment, and used as reference points without further hypothesis testing.

- Knowledge of the environment may be exploited. For example, a target in close proximity to a number of stationary targets is more likely to be stationary, and a target that is probably a part of some object already identified as stationary is also probably stationary.

A Formal Investigation of System Stability.

Finally, a formal investigation of the performance of algorithm A and of the rotation-compensated system described in this chapter would provide a rigorous assessment of the velocity extraction and noise reduction properties of the motion resolving system.

Chapter 6. Sonar Interpretation in Three Dimensions.

As they stand, the algorithms described in this dissertation are suitable for robotic vehicles operating in a two-dimensional context. For a sonar interpreter system working in a three-dimensional world certain extensions and modifications are necessary. These changes are the subject of the first part of this chapter. Also discussed here are methods for utilising external information to improve the performance of the Viewpoint Registration module.

A proficient three-dimensional sonar interpreter requires a Viewpoint Registration sub-system able to handle full three-dimensional motion, both translational and rotational. The difficulties involved in extending the motion resolution algorithms into three dimensions, as the basis of the Viewpoint Registration sub-system, encompass both scientific and implementation issues. The former concern the changes to the velocity extraction processing; the latter relate to the strong time and storage constraints that apply to any real-time interpreter implementation.

6.1 Velocity Extraction.

The major extensions necessary for three-dimensional operation of the motion resolution system involve the velocity extraction algorithms. A rigid body moving in three-dimensional space has six degrees of freedom, three translational and three rotational; in two-dimensional space it has two translational and one rotational degree of freedom. Thus the observer velocity extraction must accommodate one additional linear and two extra rotational degrees of freedom.

An extra degree of translational motion freedom is easily accommodated by the linear motion resolving system described in Chapter four. The two rotational degrees of freedom, however, require substantial changes to the angular velocity estimation scheme.

6.1.1 Angular Velocity in Three Dimensions.

In three-dimensional space the angular velocity of the observer is a three component vector, one component for each degree of rotational freedom. This contrasts with the single component vector of the two-dimensional case. Thus the extended angular velocity estimation must now find both the magnitude and direction of the vector where previously its magnitude was sufficient.

The problem is further complicated by the inobservability of the angular velocity vector. The rotational component of reflected motion is represented by the cross-product term $\underline{\omega}\times\underline{r}$ for angular velocity $\underline{\omega}$ and relative position $\underline{r}$. This vector is normal to the position vector $\underline{r}$. Thus the rotational reflected motion at any given point is constrained to lie in a plane and so has two degrees of

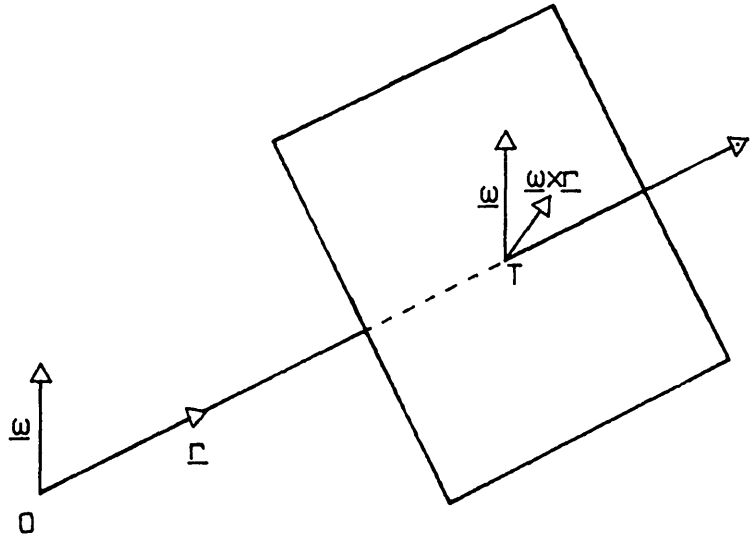


Figure 6.1. Rotational Components of Reflected Motion.

freedom (see Figure 6.1). It follows that an estimate of the angular velocity based on this reflected motion can recover only two independent components of the vector $\underline{\omega}$.

In three dimensions the identity (5.4) used in Chapter five to estimate angular velocity is invalid. Instead,

$$(\underline{r}.\underline{r})\underline{\omega} = \underline{r}\times(\underline{\omega}\times\underline{r}) + (\underline{r}.\underline{\omega})\underline{r}. \qquad (6.1)$$

The extra term, compared to (5.4), provides the component of $\underline{\omega}$ along the position vector of the point $\underline{r}$ at which computation is done. Thus the value of $\underline{\omega}$ estimated using equation (5.4) has a radial component which is artificially zero.

The angular velocity estimation problem in three dimensions is thus:

> Given a number of partial estimates of the angular velocity of the observer, $\underline{\omega}_0$, combine these two component estimates into a composite three component estimate of $\underline{\omega}_0$.

The solution to this problem is effected by an Information Averaging Filter in the system kernel. Each activated channel supplies an estimate of the projected angular velocity appropriate to its target's position. Instead of supplying an estimated error covariance for the partial estimate, the channel provides an inverse error covariance matrix adjusted so that the inverse variance along its radial direction is zero. The zero inverse variance (corresponding informally to an infinite error variance) means that no information from the radial component of the partial estimate is incorporated into the composite estimate. The artificial zero radial component of the channel's angular velocity is therefore suppressed in the combination, and equation (5.5) may be used as it stands (with the special values of Ω_j^{-1}).

Although error matrices for the partial estimates supplied by the channels do not exist (their inverses are not invertible) the Information Averaging Filter will function correctly provided that the matrix inversion implicit in its computation can be performed. The matrix to be inverted is the composite inverse error covariance, which is just the sum of the individual inverse error covariances, and will be invertible provided that at least two partial estimates are provided from channels whose target positions are linearly independent. If the Information Averaging Filter computation fails,

there is insufficient information to estimate the observer angular velocity completely.

6.1.2 Special Treatment of the Radial Velocity Component.

It was shown above that the reflected observer rotation component of the apparent motion field at any target position is confined to a plane normal to the position vector of that target with respect to the observer (see Figure 6.1). It follows that the radial apparent motion of the target contains no reflected rotational component and is entirely due to the target proper motion and observer translational velocity.

Each target radial velocity, corrected for target proper motion, constitutes a sample of the radial component of the observer's reflected translational motion field. Recalling Chapter four, that field is position independent and equal to the negative of the observer's translational velocity. Thus each target radial velocity is a measurement of the projection of the observer linear velocity onto the target radial direction and as such provides information about the observer linear velocity.

The radial velocity estimates can be used to prepare an observer linear velocity estimate in a manner analogous to the computation of the observer angular velocity estimate described above. In this case the radial velocity estimates are partial estimates of the observer linear velocity. They have only one degree of freedom, because they are constrained to lie along the target radial direction. As before, the estimation is accomplished by setting to zero the inverse variance of the estimate in those directions where it carries no information and allowing an Information Averaging Filter to compute an optimal composite based on the information actually present. In order to estimate the three degree of freedom composite from single dimensional inputs, at least three linearly independent radial velocities are needed.

The estimation of observer linear velocity using equation (5.3) already includes this special estimation from radial components. When the reflected motion estimates are corrected for the expected rotational component of observer reflected motion, the error variances in the angular directions are increased to allow for the correction term; the radial variance remains unchanged. Thus the radial velocity components are included in the estimation used in Chapter five exactly as they are in the radial estimation above, the difference being that the angular components are also included.

The advantages of estimating the linear velocity using radial components rather than the complete velocity (which contains more information) is that the former may be done without knowledge of the observer angular velocity. Thus it may be used to update the observer linear velocity estimate before a new angular velocity estimate is calculated and so improve the angular estimation accuracy. Once a new angular velocity estimate is available the information in the angular components of the reflected motion at each target can be incorporated into the observer linear velocity estimate.

6.1.3 Integration of the Observer Angular Velocity.

The orientation of the observer-relative measurement frame in the three-dimensional situation is specified by three orientation angles, one for each degree of rotational freedom of the observer. The number of angles corresponds to the number of components of angular velocity, but unlike the two-dimensional case the orientation angles are not in general simple componentwise integrals of the components of angular velocity.

As in the two-dimensional case discussed in Chapter five, observer rotation causes dynamic non-linearity in the tracking filters unless the tracking coordinate system is stabilised with respect to the viewpoint independent reference frame. The strategy for dealing with this difficulty is to compensate events obtained from the Sonar Device Interface using the current estimates of the observer orientation angles. The angular velocity estimated by the velocity extraction

algorithm is then the differential angular velocity of the orientation compensated reference frame.

The main difficulty here is the integration of the angular velocity to provide the coordinate transformation between the observer-relative measurement frame and the orientation-compensated tracking coordinate system. This problem is a well known preliminary in the study of the dynamics of rigid bodies in three dimensions, and various efficient methods for tackling it exist (for example, see Chapter two of (Wittenburg, 1977)).

6.1.4 Implementation Issues.

In view of the strong real-time constraint on the sonar interpretation problem under consideration, with its consequent limitations on the storage and time available to an interpreter, the efficient implementation of Viewpoint Registration processing is an important factor in system performance. It is likely to be a major system bottle-neck, since all data input to the interpreter must pass through the Viewpoint Registration sub-system to be motion-compensated. In this section the question of the choice of tracking coordinate system will be reviewed, focusing attention on the trade-off between theoretical desirability and potential efficiency of implementation.

The choice of Cartesian coordinates for linear motion resolution is a natural one. Linear object and target motions are transformed into linear state transitions and an optimal filtering scheme can be constructed. When observer orientation changes must also be estimated the use of an orientation-compensated frame of reference allows the linear motion resolution scheme to operate, with small corrections for differential dynamic non-linearity, as the core of a complete two or three-dimensional system.

In the case of an orientation-compensated tracking reference frame for the three-dimensional system Cartesian coordinates are no longer an obvious choice. The corrections introduced during orientation-compensation couple together the angular measurement errors with the

result that, as in the two-dimensional system, coupled Kalman filters are necessary. However, the coupled three-dimensional filter requires 27 error covariance components to be manipulated during prediction and estimation, where the two-dimensional filter required only ten. The time complexity of the filter is also $O(n^2)$ for an efficient implementation, where n is the number of components in the filter state vector. The coefficient of n^2 is substantial when the calculation of probabilities required by segmentation processing is included in the cost.

In their favour, Cartesian tracking coordinates have the advantages of linearity and the fact that the viewpoint independent reference frame is in itself Cartesian. The latter point allows orientation-compensated Cartesian tracking coordinates to be identified with the viewpoint independent reference frame apart from a known translation (the observer's tracked position), thereby eliminating the need for rotational transformations between channels and kernel in the motion resolution system.

From an efficiency standpoint the most appealing choice of tracking coordinates is the observer measurement frame equipped with propriocentric spherical polar coordinates. Under such an arrangement the noise components in a measurement vector are independent and their variances are constant. The state transition noise is also uncoupled in these coordinates since it is isotropic in Cartesian coordinates, but the covariances of the angular transition noises are range-dependent. These noise properties allow the Kalman tracking filter to be decoupled into three one-dimensional filters, one of which has constant noise variances and might further be simplified (the radial one). The effect of this decoupling is to multiply the computational cost by three (there are three single dimensional filters) while reducing n from three to one; a substantial net saving will accrue. The filter storage requirement also diminishes to nine error covariance components.

Unfortunately, for the three-dimensional case when observer orientation is allowed to vary, the full advantages of the polar frame cannot always be realised. Observer roll rotates the measurement

azimuth and elevation axes with respect to the heading axes associated with the vehicle; the changing azimuth and elevation of the observer heading vary the orientation of the heading axes. Thus, if the orientation angles are changing and must be estimated there is no measurement frame common to all measurements from a given target. Previous measurements are expressed in a frame with axes rotated with respect to the current measurement frame and consequently the noise components in the tracking filter states are coupled between the angular dimensions. The result is that a coupled angular tracking filter must be used (the radial filter may remain decoupled); the total computational cost is the sum of the costs of a one and a two-dimensional filter.

The principal disadvantage of polar coordinates is their dynamic non-linearity. Although angular estimation is easy in the polar frame, linear velocity estimation is more complex and linear motion compensation, unlike orientation-compensation, is difficult to achieve. The effect of dynamic non-linearity may be a serious one -- the experiments of Chapter five showed that a 50 milliradian per second angular velocity caused severe dynamic non-linearity and the equivalent linear velocity for a target at 20 metres range is one metre per second. Thus an observer velocity of 1 metre per second (about two knots) could cause severe distortion in polar tracking filters.

The choice between efficiency and correctness in this case is one that must ultimately be settled empirically. The crucial balance is between the cost in terms of space and time of coupled Cartesian tracking filters and the severity of the disruption caused by using a dynamically non-linear coordinate system for tracking.

6.2 Incorporating External Sources of Information.

The motion resolution algorithms described in this dissertation function with no information other than that which can be obtained from event-vectors returned by the Sonar Device Interface. No direct observer data is assumed. In practice, however, some observer motion

parameters are often accessible for small marine vehicles. The vehicle control system will maintain estimates (not necessarily highly accurate ones) of the current velocity, position and orientation of the submersible. Vehicles may carry transponder triangulation navigation aids, like that used with ANGUS, or may be equipped with motion sensors such as rate gyros, or depth monitoring equipment.

In a practical application, therefore, the sonar interpreter would not have to operate entirely from sensory data obtained acoustically. There will be available to its Viewpoint Registration module a variety of sources of direct observer motion information with varying characteristics, precisions, and accuracies. The interpreter may be expected to provide improved observer motion estimates, incorporating acoustic sensory information, to assist in vehicle control or navigation systems. The work of Lane (1984) described in Chapter three constitutes an exploratory step in this direction.

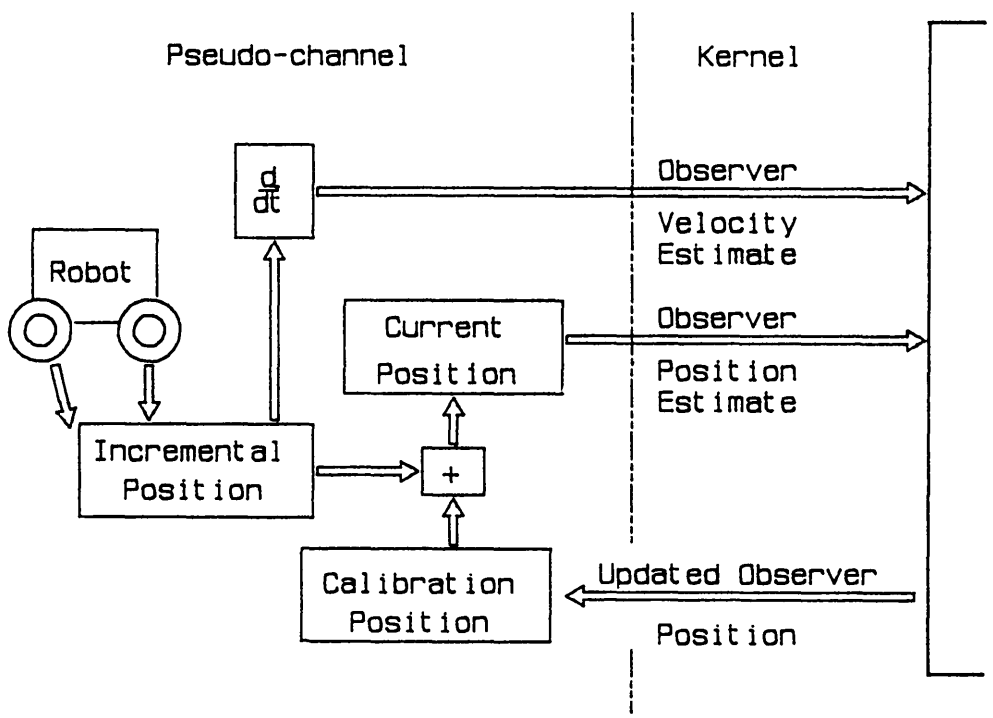


Figure 6.2. Using Direct Observer Motion Information.

Direct observer data is readily incorporated into the motion resolution systems through the Information Averaging Filters in the observer-tracking kernel. Estimates of observer state parameters, whether position, velocity or differential angular velocity, are

passed from channels to the kernel accompanied by their estimated error covariances. Provided that the same basic quantities are available from external direct sources of observer state parameter estimates, those sources may be treated exactly like channels for the purposes of input to the kernel. Treating them as channels for output purposes also would give them access to updated observer state estimates generated in the kernel.

This use of direct information sources is exemplified in Figure 6.2 where a land-based vehicle equipped with incremental shaft-encoders and an acoustic sensor is using a two-dimensional complete motion resolution system for its navigation processing. The shaft-encoders provide good incremental position accuracy but suffer from wheel-slip errors and the like in the long term. The direct observer motion information is represented by a pseudo-channel that is continually activated. The channel provides observer position estimates (and, if the transducer outputs are filtered using statistical trackers, it may provide velocity and angular velocity estimates also) and uses the observer position outputs from the kernel to recalibrate its internal position estimate relative to which the transducer data is interpreted.

In some instances it is desirable to go further and incorporate the external information sources into the kernel itself. The resulting arrangement compares directly with the terrain-aided navigation systems of section 3.2.2, where observer parameters obtained from an inertial navigation system are combined with sensory data from radar altimetry.

This style of utilisation of direct observer state data allows the incorporation of other bodies of vehicle knowledge also, for example vehicle dynamic characteristics. The kernel of each motion resolution system discussed in this thesis has assumed a linear vehicle dynamic model for simplicity; the dynamics of real vehicles is complex, coupled, and non-linear, but can be successfully modelled -- this has been done for ANGUS by Russell and Bugge (1981). Replacing the simple linear estimates of the kernel with a more sophisticated coupled model would make the system more robust and improve the

accuracy of observer tracking and velocity extraction.

Chapter 7. The Segmentation Problem.

Low level processing associated with sonar interpretation handles the construction and maintenance of a stationary global frame of reference in which the higher level processing is carried out. This task is accomplished by tracking sonar targets in the neighbourhood of the sonar transducer and using these tracks to deduce or resolve the proper motions of observer and targets. This motion resolution problem was discussed in Chapters four and five. In this chapter I consider the companion problem -- the task of matching echoes received by the sonar transducer to the targets that caused them -- which I shall term segmentation, in analogy with the visual problem of dividing an image into regions (or the general problem of dividing sensory data into perceptual units).

In recent years, the segmentation problem has increasingly been the focus of work on radar and sonar tracking systems. Considerable effort has been expended on the design of theoretically sound schemes that are also practically efficient. Therefore I make no original contribution to the solution of segmentation in this chapter. Instead, the segmentation problem will be defined, and its interrelation with motion resolution explored, in the context of a critical review of the abundant literature on radar data-correlation. The final section of this chapter collects the ideas presented into a composite suggestion for a viable three-dimensional segmentation system suitable for sonar interpretation and considers the implementation of that system for real-time operation.

7.1 The Two Faces of Segmentation.

The segmentation problem is, informally, the problem of classifying each member of a set of target sightings (events)[1] supplied by

1. In the radar references reviewed in this chapter, a different terminology is used. Raw input data from the radar is called a "return", and may be caused by targets, ground reflection, or noise. The returns are filtered to eliminate some of the clutter (often the stationary or slow-moving returns) and the resulting data are called

the sonar equipment. These sets of events arrive intermittently and represent collections of simultaneous or near-simultaneous echoes detected by the sonar device. Any event in such a set may have one of three causes: it may be an echo from a target known to the motion resolving system; it may be a sighting of a new target; or it may be a "clutter" point caused by equipment noise or by environmental "noise" such as fish or air bubbles in the water. Varying numbers of events are generated in a sonar scan and there is no guarantee that any given scan will contain an event caused by a given known target. We shall assume that events are anonymous -- that is, they carry no information other than position that might identify them.[1]

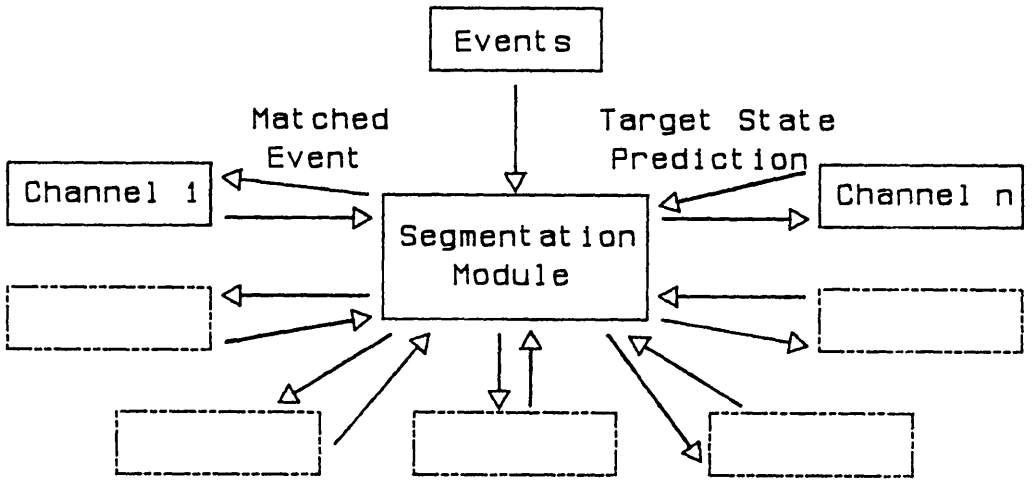


Figure 7.1. The Relationship of Segmentation and Tracking.

Segmentation is set in the context of a target tracking system (such as the motion resolving system of Chapter four, for example). The interrelationship is shown in Figure 7.1. The tracking system comprises a number of channels, each holding data associated with a

"plots". A plot is the equivalent of our "event", except that "event" does not imply any prefiltering to reduce stationary clutter as motion resolution processing is very interested in stationary targets. "Returns" correspond to our "echoes"; "tracks" are collections of associated plots, analogous with "tracks" here. Segmentation is called "data correlation" or "data association".

1. A sonar equipment might provide estimates of target surface structure from the echo spectral distribution, or the events may contain target length information. Such additional data is not a prerequisite for segmentation but it is clearly desirable that a segmentation scheme be able to use it if the sonar can supply it.

particular target. Each channel provides a prediction of the current target state (for example, its position and velocity) which is used to predict the event data expected for a sighting of that target. The input event sets are classified by the segmentation module in relation to these predictions. The result of classification is a set of matches between actual events and target channels, which is used to control the updating of the information in the channels or to initialise new channels as required.

There are two useful perspectives on the problem of segmentation: the explanation generator (EG) metaphor and the constrained relationship (CR) metaphor. Each affords insight into the underlying problem and is helpful in formalising segmentation in the context of a sonar interpreter system.

The EG metaphor describes segmentation as the task of constructing explanations for each event in the input set sequence. The explanations proposed classify events according to their three basic causes:

- the event was caused by a particular known target (one for which a tracking channel already exists);

- it was caused by a new target (for which no channel currently exists);

- it was a clutter event.

These three explanations correspond to the three basic actions taken by the Viewpoint Registration system in dealing with an event -- an existing channel may be activated, a new channel initialised, or the event ignored. This correspondence between explanation and action allows us to define the output of the segmentation module (the explanations) as a data structure describing the actions requested of the Viewpoint Registration system for each event.

In regions where several targets are close together, or in particularly noisy regions, it will often be impossible to disambiguate a number of track-event correspondences. There are several valid

explanations available to the segmentation process that cannot, on the evidence currently available, be discriminated. In such a situation, the segmentation process is able to record all the viable explanation hypotheses. If it is later possible to resolve the ambiguity using new events or higher level inference (for example, reasoning about the topology of echo sources or the presence of other objects), then the recorded explanation hypotheses are available to direct the updating of the data in the motion resolution subsystem.

The data structures generated by segmentation to describe the various explanations hypothesised from the input data will be called history trees. Each tree is associated with a hypothesised target and describes those trajectories of that target supported by evidence in the sensory input sequence. Each path from the history tree root to a leaf defines a sequence of events explained as sightings of the target, and thus a target trajectory (or track).

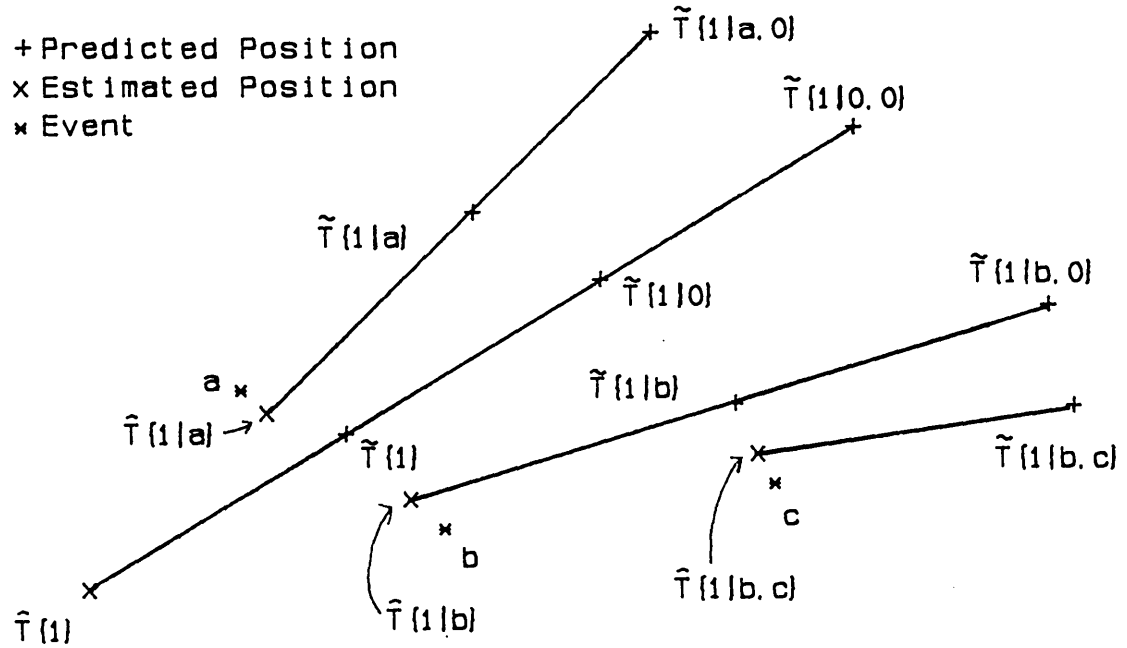


Figure 7.2. Target 1 and Several Events.

An example will clarify the situation. Suppose that there is a single target, target 1, hypothesised as travelling from the bottom left towards the top right of Figure 7.2. The current estimated position of the target is denoted by the × at the bottom left of the figure. The label $\hat{T}\{1\}$ signifies that the point is a state vector estimate for target 1 (for clarity, the figures show only target

positions).

The first event set shown in Figure 7.2 comprises two events,'a' and 'b', close to the predicted position of the target, $\tilde{T}\{1\}$.[1] Either or neither of these may be a sighting of the target. These three possibilities correspond to the three initial branches of the history tree for target 1 shown in Figure 7.3. Each node in the tree is labelled with the estimated target state for that hypothesis, and each arc is labelled with the event it explains. For simplicity we shall neglect, for the present, the possibility that either or both of the events 'a' and 'b' may be first sightings of new targets.

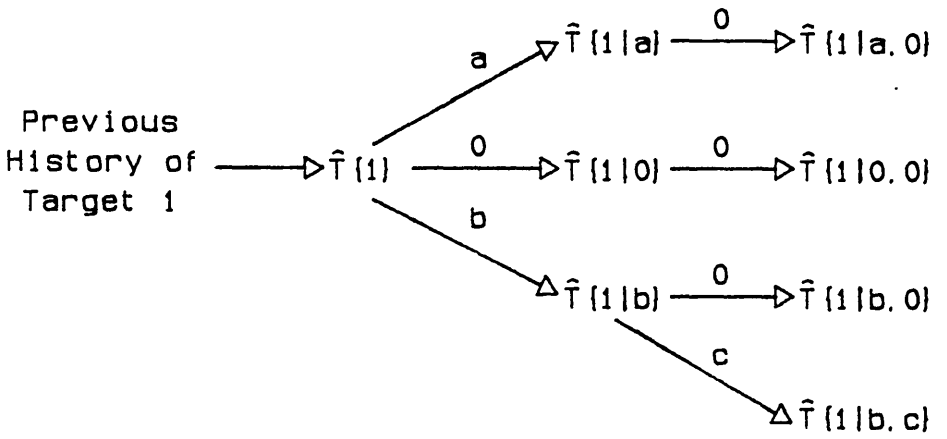


Figure 7.3. The History Tree for Target 1.

Each match between an event and the target causes a new estimate of the target state to be computed, based on the event position and the current prediction of the target state. These new estimates are $\hat{T}\{1|a\}$ for event 'a' and $\hat{T}\{1|b\}$ for event 'b' (where the event name after the vertical bar indicates that this estimate is conditional on that event being a sighting of the target). The possibility that both events are clutter also results in a target position estimate $\hat{T}\{1|0\}$ (the zero indicates that no event matched the target) which is equal to the current predicted state. Each of these estimates is used to compute new predictions of the target position when the next

1. In this example, I am using the tilde-caret convention to differentiate the predicted and estimated state vectors of the target. However, I shall also use a conditioning notation to indicate explicitly the event-track matches on which the estimates and predictions depend.

event set arrives.

The next event set consists of the single event 'c'. Suppose, again for simplicity, that event 'c' is too far from $\tilde{T}\{1|0\}$ and $\tilde{T}\{1|a\}$ to be a reasonable match for either of these tracks. Then, neglecting the possibility that it is a new target sighting, it is either a clutter point or a match for the track that generated prediction $\tilde{T}\{1|b\}$. The first two possibilities generate new estimates $\hat{T}\{1|0,0\}$ and $\hat{T}\{1|a,0\}$ equal to the respective predictions. The last two possibilities generate estimates $\hat{T}\{1|b,0\}$ and $\hat{T}\{1|b,c\}$. These four current state estimates for target 1 are shown as four leaves on the history tree in Figure 7.3. Thus there are now four hypothesised tracks for target 1, corresponding to the four paths from the tree root to the leaves.

When the possibility of an event being the first sighting of a new target is admitted, each event causes the construction of a new history tree describing that target. Similarly, when multiple targets are being tracked, it may be possible to explain an event in several ways. One event can therefore label arcs in several history trees.

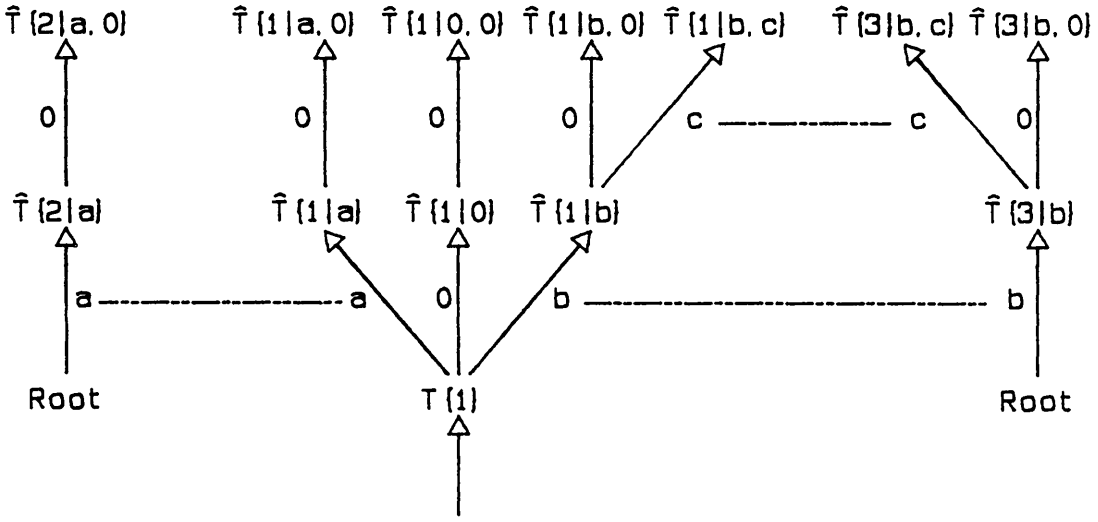


Figure 7.4. Cross-linking Explanations between History Trees.

The multiple explanations of any single event are mutually exclusive, since an event may have only one cause. The arcs representing exclusive explanations may be cross-linked, independently of the history trees. To illustrate this, suppose now that the possibility of events 'a' and 'b' being the first sightings of

new targets is taken into account. Each of these events initialises a new target state estimate and starts a new history tree. These are illustrated in Figure 7.4 as $\hat{T}\{2|a\}$ and $\hat{T}\{3|b\}$, and the sets of cross-linked arcs are joined with a dashed line. Each cross-linked arc set then represents the possible (exclusive) explanations of the given event.

The second perspective on segmentation, the CR metaphor, follows naturally from the observation that cross-linked explanations are mutually exclusive. Certain combinations of explanations are incompatible. The CR metaphor views the goal of the segmentation process as the construction of a self-consistent constrained mapping between sets of events and sets of tracks. It formalises the principle that the set of explanations of an input event set, and indeed the set of explanations described by any history tree, is constrained by the nature of the sonar imaging process and by the properties of the world in which the interpreter operates.

Incompatible sets of explanations correspond to illegal maps in the CR metaphor. The constraints that a legal mapping (a legal combination of explanations) must satisfy are enumerated below.

(1) An input event may match at most one target track.

(2) A track may match at most one event in a given input set.

These constraints correspond to the exclusivity of cross-linked arcs, and the exclusivity of arcs leaving the same node in a history tree. They embody the principle of causality. Each non-noise event in an input set has one, and only one, explanation. A target track may continue in one, and only one, way (note that tracks which merge are dealt with by combining cross-linked branches of the history tree(s) involved).

(3) A track may only match events that are physically close to its predicted target position.

This constraint is the property of locality -- the echoes generated by sources are physically near to the sources themselves.

(4) Any or all of the events in an input set may be clutter points or new target detections.

There is no way of distinguishing a new target sighting from a clutter point in the CR metaphor, since the mapping constructed is between existing tracks and current events.

In the CR metaphor, there are two candidates for the event set used in the construction of the map. It may be either the current set of events obtained from the sensor or the union of all the event sets received so far. The constrained mapping construction for the second possibility with the example of Figure 7.2 used above is as follows.

Tracks	Matched Events			Matched Events		
	a	b	c	a	b	c
$\hat{T}\{1\|0,0\}$	0	0	0	0	0	0
$\hat{T}\{1\|a,0\}$	0	0	0	1	0	0
$\hat{T}\{1\|b,0\}$	0	0	0	0	0	0
$\hat{T}\{1\|b,c\}$	0	0	0	0	0	0
Clutter	1	1	1	0	1	1
	Map 1			Map 2		
	a	b	c	a	b	c
$\hat{T}\{1\|0,0\}$	0	0	0	0	0	0
$\hat{T}\{1\|a,0\}$	0	0	0	0	0	0
$\hat{T}\{1\|b,0\}$	0	1	0	0	0	0
$\hat{T}\{1\|b,c\}$	0	0	0	0	1	1
Clutter	1	0	1	1	0	0
	Map 3			Map 4		

Table 7.1. Legal Mappings for the Example in Figure 7.2.

First, the set of candidate tracks is enumerated. These are generated from the set of all sequences of events using the locality constraint (3). The four consistent mappings between events {a,b,c}

and this track set are given in Table 7.1. (The tracks are labelled with the name of their leaf in the history tree.) The four possible mappings are exclusive and (neglecting new target creation) exhaustive, and correspond to the four possible tracks represented in the history tree of Figure 7.3.

In either metaphor, the description constructed by segmentation contains a number of exclusive sets of hypotheses that together exhaust the set of all possible explanations of the input event sets. The number of hypotheses to be considered grows rapidly with time. Practical segmentation algorithms calculate confidence levels for these hypotheses and implement resource management schemes (for example, eliminating all sufficiently unlikely hypotheses). The next three sections review practical algorithms applied to track updating, track initiation, and target manoeuvre handling.

7.2 The History of Segmentation.

Target tracking techniques for sonar and radar sensors have aroused a great deal of interest over the last twenty years, both in military and civilian circles. During that time various attempts have been made to treat the segmentation problem rigorously and to derive statistically optimal algorithms for its solution. A considerable effort has been expended in developing systems that are theoretically sound and achieve a high standard of tracking accuracy, robustness and real-time performance. Most of this effort has focussed on the track updating problem, variously referred to as "data association" or "report-to-track correlation" in the literature, both in attempting to resolve or overcome the ambiguity of event-track matching in cluttered or multiple target environments and in accommodating the malicious, evasive, or accidental changes of velocity, called "manoeuvres", executed by targets during tracking.

The standard tracking techniques are now the α-β filter and the (possibly extended) Kalman filter, discussed in Chapters three and four. In this review I shall attend principally to the algorithms which, using information held in the filter states, control the

disposition of events among these filters and the creation of new tracking channels. I shall deal separately, as far as is possible, with the three major areas of current research: correlation in dense multitarget environments; handling target manoeuvring; and track initiation. There is considerable overlap, for a successful automatic tracking algorithm must have some competence in all three areas.

7.2.1 Data Association Ambiguity.

Brief allusion was made in Chapter four to the intimate connection between the updating of track state and error covariance estimates and the process of segmentation that provides the data used for updating. It was pointed out as early as 1964 by Sittler (Sittler, 1964) that the position measurements supplied to a tracking filter by its data association algorithm may in fact be incorrect. This uncertainty contributes to the filter state error and a tracker that neglects the possibility of segmentation error will be optimistic in its assessment of its own performance.

A basic solution to the problem of handling data association ambiguity was suggested by Sittler and remains popular today. Wherever it was necessary to incorporate uncertain measurements into a track, he retained all the plausible hypotheses by splitting the track as many times as necessary and using each member of the track sheaf to follow one hypothesis. The likelihood function was computed for each hypothetical trajectory and sufficiently unlikely hypotheses were dropped. Sittler's work was done before Kalman filtering became standard in tracking systems. His ideas have been extended into the newer framework; this and the modern Bayesian approach are described by Bar-Shalom in his survey paper (Bar-Shalom, 1978).

7.2.2 Likelihood Methods.

Following Bar-Shalom I shall divide the approaches to segmentation error handling into two groups: the (non-Bayesian) likelihood methods; and the Bayesian probabilistic association techniques.

Within each classification there are algorithms that involve track splitting (i.e. simultaneous consideration of a number of alternative hypotheses) and those that maintain at most one current hypothesis per target. Additionally an algorithm may be batch-orientated or sequential (recursive) in structure. The former category require a collection of input data over a period of time on which to operate whereas the latter use only the current inputs and previously calculated results.

A similar approach to Sittler's, but within the Kalman filtering framework, was developed by Smith and Buechler (1975) for radar target tracking. They recursively compute the log-likelihood function for each track hypothesis in terms of the Kalman filter innovation probability distribution at each stage of the updating implied by the hypothesis. To limit storage requirements that otherwise would grow exponentially with time as measurements arrive, the set of hypotheses under consideration undergoes a threefold pruning: the log-likelihoods of individual track-event matches are thresholded to eliminate unlikely tracks; track hypotheses are eliminated when their support (i.e. cumulative log-likelihood) falls below a given value; and sets of physically close trajectories are pruned so that only the best supported (most likely) one remains. The method may be extended to allow for uncertain detection (Smith and Buechler assumed perfect detection) and, given good target state prediction, it appears to function satisfactorily. If, however, the state prediction is poor, the plausible hypotheses multiply rapidly and saturate the available storage.

The estimates of state and error covariance are obtained from standard Kalman filters for each hypothesis. There is, however, no way of determining the probability that any hypothesis is correct -- the likelihoods are the probabilities that the measurements used in the hypothesis would arise <u>given</u> that the hypothesis is correct. No use is made of the constraints on hypotheses. Thus it appears, as Reid (1979) points out, that a target is allowed to match more than one event in an input set. The main difficulty with this technique is the combinatorial explosion of plausible hypotheses.

Morefield (1977) advanced a novel technique for dealing with this combinatorial explosion and for selecting the most likely (*a posteriori*) collection of compatible hypotheses using standard methods of integer programming. His algorithm is a batch one, that is, it uses all the data collected up to the present time in a batch, and is able to deal with false detections, uncertain target detection, and *a priori* knowledge of clutter density and detection probability if available. The number of targets must be known in advance.

The algorithm is best understood in the CR metaphor. First, a set of feasible tracks is enumerated using a simple sequential window test on the input measurements. Then a set of composite hypotheses is defined; a composite hypothesis comprises one track hypothesis from the feasible track set for each target, and a clutter hypothesis. Together the components of a composite hypothesis account for all the measurements collected so far. For each feasible track the negative log-likelihood is computed.

In terms of the example given in Figure 7.2, the feasible tracks are listed below in Table 7.2. The remainder of this table, which is similar to Table 7.1, is a matrix that specifies whether a given event is used in a particular feasible track.

Track	Uses Event		
	a	b	c
$\hat{T}\{1\|0,0\}$	0	0	0
$\hat{T}\{1\|a,0\}$	1	0	0
$\hat{T}\{1\|b,0\}$	0	1	0
$\hat{T}\{1\|b,c\}$	0	1	1

Table 7.2. An Example of a Feasible Track Set.

Morefield then casts the consistency enforcement problem into a 0-1 integer programming of the set partitioning or the set packing class. To do this, he represents a typical member of the composite hypothesis set as a binary vector $\underline{\rho}$ where element i indicates whether feasible track number i is a part of the composite hypothesis. The

total likelihood of any composite hypothesis is then $\underline{\lambda}^T\underline{\rho}$ where $\underline{\lambda}$ is the vector of feasible track log-likelihoods. Then a binary matrix A is constructed such that a_{ij} indicates whether measurement i is required by feasible track j. (The binary matrix in Table 7.2 is the matrix A for the example above.) With this definition the constraints 1-4 (page 205) can be expressed as the equation $A\underline{\rho} \leq \underline{1}$ where $\underline{1}$ is a vector of ones of suitable dimension. Choosing the most likely self-consistent composite hypothesis amounts to minimising $\underline{\lambda}^T\underline{\rho}$ where $A\underline{\rho} \leq \underline{1}$ and $\underline{\rho}$ is a binary vector. Morefield uses the techniques of 0-1 integer programming to evaluate $\underline{\rho}$ without requiring the enumeration of the composite hypothesis set, which may be very large. Once $\underline{\rho}$ is determined state estimates and error covariances are computed, using the standard Kalman filter equations, for each track.

This approach, while attractive because of the implicit enumeration of the composite hypothesis set, is unsatisfactory in practise because it requires a priori knowledge of the true number of targets present and because of its batch-oriented structure. It is theoretically unsatisfactory because the state and error estimates computed are conditioned upon the truth of the chosen (most likely) composite hypothesis. Neither method considered so far attempts to allow for the possibility that its hypothesis is incorrect when updating the filters.

A system following the track splitting approach of Sittler and of Smith and Buechler has been described by Miller (1981). This system is designed for single target tracking, and handles target manoeuvres using a likelihood computation scheme based on Smith and Buechler's. Miller calls each track hypothesis an "option". Storage control (the algorithm implies exponential growth of store) is achieved by eliminating track hypotheses using a per-scan threshold and by discarding the lowest likelihood hypothesis whenever the number of options considered would exceed a preset maximum. The significant difference in our context is that Miller attempts to include the uncertainty of segmentation by using a weighted mean of the option state estimates as a composite estimate of the target state. The weights used are the relative likelihoods. He does not compute error covariances for this estimate because the system uses a constant gain (and therefore

sub-optimal) α-β tracking filter. His approach is thus a compromise between the likelihood method and the Bayesian methods to be described next.

In a performance analysis of the system, Miller shows that an eight option system is adequate for tracking a strongly manoeuvring target; with merging of similar options and pruning techniques storage use is economical. He also points out the important correlation between tracking failure, where the target is lost, and saturation of the available storage.

7.2.3 Bayesian Techniques.

The likelihood methods already described are sub-optimal but simple to compute and are able to handle track initiation as well as track updating. The Bayesian techniques presented here are computationally more intensive, theoretically satisfactory and potentially optimal, but most of the algorithms developed assume that track initiation has been done -- they address only the track updating aspect of segmentation. The fundamental difference between the Bayesian and likelihood approaches is most clearly seen using the EG metaphor. The likelihood approach assumes that a particular set of hypothetical explanations is correct and computes the probability that the measurements thus explained actually arise -- that is, it computes measurement occurrence probability given the correctness of an explanation. Bayesian methods calculate the probability that a particular explanation or set of explanations is the true one, given the sets of events that have been collected.

The likelihood approach calculates the conditional measurement occurrence probability for each feasible explanation set, and the set (or sets) with highest probability is the explanation chosen. In the Bayesian framework, once the probability of correctness of each explanation is known, the state and error covariances of several explanations may be combined in a statistically optimal manner.

Early work in the Bayesian framework was done by Singer, Sea, and Stein (Singer and Sea, 1973; Singer and Stein, 1971). They developed algorithms for tracking several targets in a realistic cluttered environment. They proposed associating each target with the event closest to the target's predicted position -- its "nearest neighbour" -- and adjusted the Kalman filter statistics to account for possible segmentation error using the a priori probability that the chosen association was correct.

Singer and Sea (1973) developed this approach. The optimal a priori filter has expanding storage requirements because of the multiple hypotheses to be considered. In contrast they presented a suboptimal a priori filter with a constant storage requirement by considering the three possible hypotheses at each data arrival time -- that the target was not detected; that it was detected but was not the nearest neighbour; and that the association chosen was correct. At each stage the hypotheses were combined based on their a priori probabilities of correctness.

Jaffer and Bar-Shalom (1972) proposed an a posteriori version of that nearest neighbour filter. Their algorithm computed the a posteriori probabilities of possible explanations based upon the relative position of target prediction and nearest neighbour. Later in that year Bar-Shalom and Jaffer (1972) pointed out that all possible data associations with a given target should be considered, rather than just the nearest neighbour. A final version of the algorithm proposed there, the Probabilistic Data Association (PDA) technique, was given later by Bar-Shalom and Tse (1975).

The PDA method assumes a single target in clutter. The track must already be established. Clutter is modelled as independent random measurements identically distributed uniformly throughout the observation volume; the number of clutter points observed at any time may be modelled as a Poisson process, and the clutter density and target detection probability are assumed to be known a priori. After receiving an event set the PDA method computes a posteriori probabilities for each possible explanation of the input data -- for m current input measurements there are m+1 explanations, viz. all

measurements incorrect, and all but measurement i are incorrect for each i between 1 and m. The state estimate generated using the current data is then the conditional expectation (that is, the average weighted by probability) of the states associated with the explanations. The covariance matrix is also augmented to account for the use of uncertain data.

The PDA algorithm allows no track splitting and therefore it requires constant storage only slightly in excess of that needed by the Kalman filter. To reduce the computational load the "all neighbours" ambiguity is restrained by neglecting those events that lie outside a "validation gate" centred on the target predicted position. The gate is a confidence hyperellipsoid defined by a threshold on the manoeuvre distance (the distance, weighted by the innovation covariance, between the event and the predicted target location) of an event-target match. The size of the gate determines the probability that the true target return, if detected, falls within the gate.

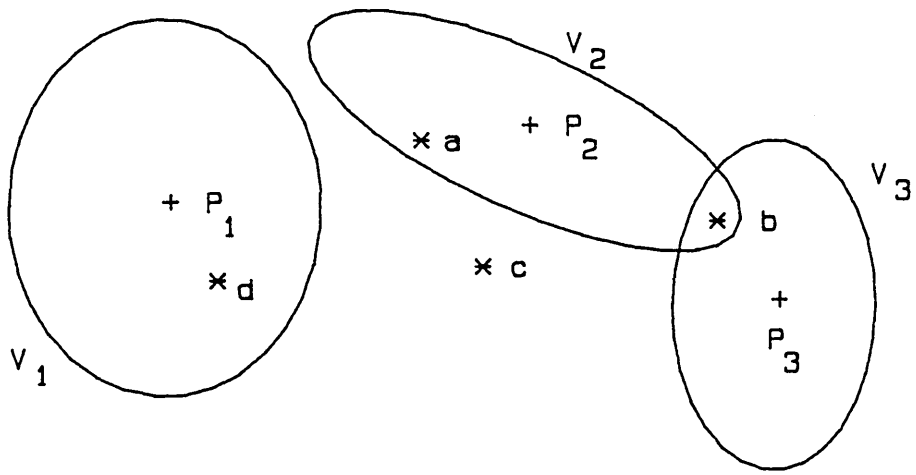


Figure 7.5. An Example of Validation Gates.

A two-dimensional example of the use of gates is shown in Figure 7.5. Suppose that there are three targets, at predicted positions P_1, P_2 and P_3, and that four events ('a'-'d') have been detected at the positions shown. The validation gates, whose size and shape depend on the prediction and measurement error variances for the targets, are the ellipses V_1, V_2 and V_3. In the situation shown, the

only match for target P_1 is the event 'd', since all the other events are outside the gate V_1. Similarly event 'c' may not match any target, and 'b' may match either P_2 or P_3. Targets P_2 and P_3, whose validation gates are "connected" by the event (Reid, 1979; Bar-Shalom, 1974). P1 also constitutes a (trivial) cluster.[1] When only a single event falls inside the gate the PDA filter reduces to an ordinary Kalman filter (with an allowance being made for the possibility that the detected event is really clutter).

The PDA filter, since it does not use track splitting, is a sub-optimal filter. Concurrently with the PDA development, Singer, Sea and Housewright (1974) published details of an optimal *a posteriori* filter using the "all neighbours" Bayesian approach (i.e., unlike the PDA filter, they do not restrict event-track matches using validation gates). Instead of considering just the possible explanations of the current set of measurements, the optimal algorithm constructs all the possible explanations of the data accumulated so far. This may be done sequentially and needs exponential storage to hold the history. The *a posteriori* probability of correctness is then calculated for each branch of the history tree. The state estimate generated by the algorithm is (like the PDA method) the conditional expectation of the state estimates for the various track hypotheses -- in this case the whole of the history tree. The error covariance is the sum of the conditional expectation of the individual track-hypothesis error estimates and a term compensating for the uncertainty in the state estimate caused by the segmentation ambiguity. Since all the data computed before the current set of measurements is available for use, the algorithm operates recursively.

Of course, the optimal filter requires unbounded storage and computational resources. A practical implementation requiring constant resources may be defined at the expense of optimality. Singer *et al* define the class of N-scan-storage filters (where N is a fixed parameter) as a realistic possibility. An N scan (or N backscan) filter accumulates history branches until there are N+1 sets of input data

1. Notice that events such as 'b', which are responsible for the clustering of targets, have several possible explanations, so the history trees of targets in a cluster are cross-linked.

available. After completing that cycle, the branches which are identical from scan 2 to N+1 inclusive are combined into a new, single, branch. This procedure is repeated after every subsequent filter cycle -- it amounts to neglecting history more than N scans old. A zero-backscan filter contains no history; all the possible explanations are recombined at every scan. The PDA algorithm is an example of such a filter. Alternative practical sub-optimal implementations eliminate all but the L nearest neighbours of a predicted target position, or vary their storage requirements dynamically according to demand.

In a theoretical and experimental performance assessment of the filters, Singer et al. conclude that: the one and two backscan filters provide almost optimal performance; that the "all neighbours" a posteriori filters perform at least twice as well as their corresponding a priori versions; and that the optimal a posteriori filter under high clutter conditions estimates its own performance as about 30 times worse than the standard Kalman filter.

A graph of track loss[1] probability against normalised clutter density (normalised with respect to measurement error) is also given in their paper.

The a posteriori filters just described are designed to track a single target in clutter. They assume that the track is already initiated. These ideas have been extended to work with an arbitrary number of targets in the Joint Probabilistic Data Association (JPDA) technique of Fortmann, Bar-Shalom and Scheffé (1983). The algorithm described is an improvement on the PDA filter; it is a zero backscan filter in which account is taken of all the possible sources of each event in the input set.

The major difficulty besetting PDA in a multitarget environment is the non-random nature of the interfering measurements. Whereas noise may be adequately represented as independent Poisson clutter points, that model breaks down dramatically for the consistent detection of

1. Singer et al define "track loss" as that condition where, owing to large state errors, a track is consistently updated using clutter points.

events from interfering targets. To deal with this the JPDA method computes the probabilities of correctness of the possible explanations of each event jointly across the set of events and targets; the PDA approach computes these probabilities independently for each target, assuming that all events not associated with that target are false.

In the JPDA algorithm, targets are grouped into clusters -- sets of targets whose validation gates are "connected" by events lying in their intersections (see Figure 7.1) (Reid, 1979; Bar-Shalom, 1974) -- and each cluster is processed independently as follows. The set of feasible hypotheses is constructed from the set of all possible event-track matches using the consistency constraints 1-4 (page 205) described in section 7.1. The joint a posteriori probabilities of those hypotheses are computed using Bayes rule and suitable assumptions about the nature of clutter and measurement noise. The probability that a particular event matches a given track may then be computed by summing the probabilities of those joint hypotheses for which this is true. Finally, the state estimate for each target is computed using the conditional expectation over the hypothesis set, as in the PDA filter.

A simulation run employing a pair of heavily interfering targets showed that the JPDA method performed satisfactorily; the PDA method in the same situation lost track of the targets and both tracks locked together in a compromise track.

As it stands, the JPDA algorithm is sub-optimal, since it is a zero backscan filter. However, Fortmann et al. argue that including past scans in the decision process would greatly complicate both the filter and the manoeuvre detection system built on top for very little gain in real performance.

The final algorithm considered in this section is that of Reid (1979). Whereas the Bayesian methods already described assume that track initiation has been done beforehand, Reid's algorithm handles track initiation in the same framework as track updating. To my knowledge it is the only algorithm in this class that does so. The track initiation processing will be described in the next section.

Reid's method is recursive; it can handle multiple targets, false detections, and clutter points. Its basis is closer to the EG metaphor than other Bayesian techniques: rather than considering which measurements might have originated from a given target, it examines the possible explanations of each measurement in turn. This view requires that it be possible to construct target state data from a single measurement. A history tree of explanations is recursively constructed and the a posteriori probability that each branch of the tree is correct is computed using Bayes rule. This is a similar Bayesian technique to the track-splitting relative likelihood approach, except that Reid's algorithm computes probabilities rather than relative likelihoods for each hypothesis (tree branch). The combinatorial explosion of hypotheses is controlled by tree pruning and tree binding. The former operation eliminates all history branches with sufficiently small probability; the latter merges branches with sufficiently similar effects, and computes a composite state and covariance estimate from the individual branch estimates.

The filter performs adequately over a range of clutter densities and branch elimination thresholds. Reid presents results showing that the filter error estimates agree well with the actual position errors over a wide range of performance parameters. The filter correctly tracks 75% of targets under the worst conditions (high clutter density and most rigorous tree pruning) and about 93% at best conditions. Increasing the new target creation probability improves these to 91% and 99% respectively, but the false track rate under bad conditions remains relatively constant.

Reid's algorithm processes targets in clusters -- a cluster being defined similarly to those in the JPDA method, except that each cluster also includes its bundle of cross-linked history trees describing the hypotheses associated with those targets and measurements. Clusters are amalgamated when a measurement connects the validation gate of a target in one with that of a target in the other, (this is called "collision") and they are decomposed if the hypothesis reduction techniques eliminate all connection between them. This will be discussed further in the next section. The cluster-processing organisation allows full advantage to be taken of concurrent processing

equipment.

7.3 Track Initiation.

In comparison with the companion problem of track maintenance, the track initiation problem has received relatively little attention. There are a number of reasons for this. Systems for tracking are often intended for air traffic control or military surveillance operation where it is not necessary for the system to work fully automatically; target selection and track initiation may be done by a human operator (e.g. in (Fortmann, Bar-Shalom and Scheffé, 1983; Symons 1982)). Also, the track splitting techniques used for track maintenance are in principle extensible to track initiation, so that to some extent the former problem has subsumed the latter in the literature.

The track initiation problem is intrinsically more complex than track maintenance: there is very little *a priori* information to direct the segmentation choices, so the potential for combinatorial explosion is greater. To form an estimate of target position and velocity requires two target sightings, and in a noisy environment the number of possible pairs of events, even allowing for restrictions on maximum target velocities, may be large.

There have been two major directions of approach to automatic track initiation. In this section we shall review an example of each class: the likelihood methods used by Holmes (1977) and Tuncliffe (1977), and the Bayesian methods of Reid (1979). It has also been noted by Fortmann *et al.* (1983) that Morefield's 0-1 integer programming method, with its implicit enumeration of the set of mutually consistent hypotheses (Morefield, 1977), would be suitable for selecting the most likely set of tracks to initiate given a small batch of noisy data describing a number of new targets.

The tracking system described by Holmes (1977) and Tuncliffe (1977) handles track formation and updating using a likelihood[1]

1. The meaning of "likelihood" here differs slightly from the previ-

framework. The radar-driven system employs stationary plot prefiltering to eliminate all the stationary and slow-moving targets that are of no interest. Once known tracks have been updated, unexplained events are candidates for the track formation processing. Pairs of candidate events from different scans are selected, position and velocity estimates for the postulated target are computed, and the embryonic tracks are screened for unrealistic target velocity.

Any track that passes the screening is provisionally accepted by the system and is updated on subsequent scans. At each updating the log-ratio of the probability that the track is valid to the chance that it is false is computed using a staircase integrator -- the computation amounts to incrementing or decrementing the integral appropriately depending on the event currently being incorporated into the track. If the integral falls below a preset value the track is eliminated as implausible; when the integral exceeds a threshold the track is accepted as confirming the existence of the target. These thresholds are chosen for a desired constant rate of rejecting valid tracks and accepting false ones, respectively.

The method is fast and effective, but it makes no attempt to handle the statistical problems of segmentation uncertainty. The algorithm is not really a track-splitting one; an event used by the updating section of the algorithm is 'consumed' by that processing and the probability that a track is correct is never computed. It does not allow for multiple target interference, since the matching done is essentially nearest-neighbour.

Reid's algorithm, on the other hand, treats track formation and maintenance using an integrated Bayesian track-splitting framework. As mentioned above, the algorithm views segmentation as the explanation of events and maintains an explicit record of the explanation hypotheses (a set of cross-linked history trees) associated with each target cluster. In this context, track initiation requires an extra hypothesis per event -- the explanation of that event as a sighting of a new target. The probabilities of such hypotheses may be

---

ous section. Holmes employs Bayes rule in his computations, but the value being computed is the relative likelihood that a track is valid as opposed to invalid.

computed using the new target density parameter -- the expected number of new targets per unit volume -- and suitable distribution assumptions.

The algorithm processes targets in clusters whose validation gates are connected by events. At each stage of processing, hypotheses of new targets are introduced for each event in a cluster and the probabilities of various hypothetical explanations are computed as described in the track updating section. Then the tree pruning and binding procedures are applied to the history tree. At this point the algorithm attempts to perform a further simplification of the hypothesis set that will result in the decomposition of the cluster and the initiation of new confirmed targets. All those events that are uniquely explained (once negligible possibilities are dropped) are removed from the cluster history tree and used to create new clusters containing the targets that explained them (events explained as clutter do not set up new clusters). This process decomposes the original cluster in preparation for the next cycle of processing. If a target so transferred is a "tentative target" then it is marked confirmed; the criterion for new target acceptance is thus that the new target is necessary to explain a particular event. After each processing cycle an updated estimate of the new target density is made for use during the following cycle.

Reid presents results that illustrate the track initiation properties of his algorithm. In one case the algorithm successfully confirmed a single unknown target after five observations of that target, in the absence of other targets and noise (the algorithm still expected noise and missing data in its input, however). In the second example, the filter correctly deduced a fifth target, given four a priori targets and noisy data, after nine sets of events.

The parameter relevant to this aspect of the filter's performance is the relationship between new target density and false event or clutter density. The size of the former determines the algorithm's readiness to postulate new targets to explain events; the latter controls the willingness with which data is ignored by the system. The larger the new target density, the greater the number of false

tracking hypotheses the system will entertain; if the density is too small then true targets will not be tracked.

The major problem to be expected with track-splitting algorithms such as Reid's is one of resources. The basic method requires an expanding store and computational power in order to process the exponentially growing history tree; allowing extra explanation hypotheses aggravates this. In order to contain the unbounded requirements of space and time the history trees must be pruned and there is a danger that the system will saturate its storage and fail.

Reid presents results of Monte Carlo tests suggesting that this is unlikely to happen. The tests run contained random numbers of targets of which 80% were known a priori, and under these circumstances the filter required between 4.4 and 16.4 hypotheses on average (the expected number of targets was five). Increasing both new target density and pruning rate (which increases the turnover of hypotheses) gave a good performance (90% of targets tracked with 7.8 hypotheses on average) under noisy conditions.

7.4 Target Manoeuvre.

Tracking systems that use model-based filters, such as the ordinary or extended Kalman filter, implicitly assume a particular sort of target behaviour -- that described by the filter's state transition model. The model chosen depends principally on application and the desired accuracy of tracking. For example, Singer and Monzingo (1971) use a specialised model for tracking goal-seeking attack vehicles (e.g. target-seeking missiles) that incorporates knowledge of the attacker's control law. On the other hand, the more general application of sonar tracking employs a less specialised model: targets are often assumed to move in straight lines with roughly constant velocity.

If the target behaviour does not correspond well to the model's assumptions, tracking performance will be poor and targets will often be lost -- the filter estimates will no longer be "locked" to the actual target. In military applications targets may be expected to

execute malicious evasive manoeuvres in attempting to confuse defensive tracking systems and this sort of unpredictable behaviour must in some way be taken into account in the target models if the system is to be robust. In less demanding applications targets may not manoeuvre maliciously, but nevertheless they occasionally behave in ways that the tracker models do not describe -- for example, a supposed constant velocity aircraft may decelerate, or may turn through 90°, under direction from air traffic control.

For the purposes of this review I shall consider a target manoeuvre to be a part of the target's trajectory that does not correspond to the simple constant velocity motion model. We also distinguish the more demanding "military" problem, in which targets manoeuvre unpredictably (although the classes of manoeuvre and the worst case manoeuvres expected can often be modelled), from the simpler "civilian" problem where manoeuvres tend to occur in cooperation with the tracker (for example, when an air traffic controller orders an aircraft course or height alteration). In the context of sonar interpretation, the distinction is between unexpected manoeuvring of individual targets and the collective motion of the local neighbourhood induced by planned motion of the robotic vehicle (we shall term this behaviour observer manoeuvre, since it corresponds to unmodelled observer behaviour).

In the case of planned observer manoeuvre it would be possible (and reasonable) for the control system to warn the sensory interpreter systems that a large scale vehicle motion was imminent, so that accurate tracking (and hence the sensory reference frame) could be maintained over the manoeuvre. The random perturbation of observer and targets from their expected model trajectories is not manoeuvre since it is encompassed by the transition model.

The definition of manoeuvre as unmodelled target behaviour immediately suggests the two main lines of attack pursued in the literature. Either the transition model is extended to encompass the class of manoeuvre envisaged (thereby effectively defining it away), or the effects of the manoeuvre are detected and the tracking system estimates are compensated for these effects.

7.4.1 Augmented Transition Models.

A simple model for the accelerated motion of manned targets was suggested by Singer in (Singer, 1971). His model incorporates the idea of consistency in manoeuvre that might be expected from manned (i.e. non-random manoeuvring) targets as a non-zero autocorrelation in the sequence of accelerations. The model has two parameters determining the autocorrelation according to the formula

$$E[\, a(t)a(t+\tau)\,] = \sigma_m^2\, e^{-\alpha|\tau|}$$

The time constant $1/\alpha$ determines the violence of the manoeuvre -- it is small for fast, sharp manoeuvres, and large for long, slow ones. The variance σ_m^2 is the variance of the distribution of the acceleration $a(t)$, which takes values in the interval $[-A_{max}, A_{max}]$ with non-zero probabilities that $a(t) = 0, \pm A_{max}$. These parameters are incorporated into a Kalman filter and increase the transition noise covariance.

This model has the advantage of being simple and easy to compute. However, using such a model degrades filter performance in non-manoeuvring sections of target trajectory because, with its inappropriately augmented transition noise, the filter is more susceptible to the errors in its input measurements.

A more sophisticated approach that avoids this difficulty uses an adaptive state estimation filter such as that described by Moose (1975). In this method the transition model is switched at random among a number of possible models using a Gauss-Markov process. The Gauss-Markov process extends the switching property of a Markov process to include a notion of time -- the time between Markov state transitions is a random variable τ, often with an exponential distribution. The possibilities allow for non-manoeuvring trajectory and various degrees of manoeuvre. Moose implements this adaptive filter as a set of Kalman filters, one for each transition model possibility, with machinery to estimate the correct state of the Gauss-Markov process; the estimation is done using Bayesian methods and the final target state and error variance are formed by combining the states of

individual filters using the Gauss-Markov process state estimate.

This technique requires more memory and computation than the simple filter since it consists of several Kalman filters operating in parallel. If the measurement and transition noise models are identical among all the various possibilities, the adaptive filter reduces to a single Kalman filter. However, in general the multiple transition models and state switching of the Gauss-Markov model mean that the filter is no longer linear and filter bias compensation must be included.

This adaptive filter approach has been extended by Kenefic (1981) by combination with the Bayesian a posteriori track splitting filter of Singer, Sea and Housewright (1974). The resulting filter, which illustrates the possibility of dealing with manoeuvring targets within the same framework as track initiation and updating, deals optimally with a single manoeuvring target in clutter. In common with the track splitting filters, it requires expanding resources.

7.4.2 Manoeuvre Detection Methods.

In many real-time applications the use of a theoretically optimal manoeuvre filter is a luxury that the tracking system cannot afford. Responding to this need, a number of manoeuvre detection and post-compensation schemes have been developed. The tracking system assumes that targets behave as described by the constant velocity trajectory model, and manoeuvre handling mechanisms of varying complexity are invoked if this assumption is recognised as false for a target at a particular time.

Manoeuvre detection for systems using Kalman filters can be implemented as prediction bias sensing. When the target behaves as the transition model expects, the filter innovation sequence is randomly distributed about a zero mean. If the target is manoeuvring, the consistent prediction errors caused by the manoeuvre appear as bias in the innovation sequence, shifting its mean. Methods of detecting this include tests on the sign (which should be an equiprobable Bernouilli random variable when there is no manoeuvre) and magnitude of

the innovation, as used by Demetry and Titus (1968); tests on the average innovation bias, as used by Tenney, Hebbert and Sandell (1977); and tests of the χ^2 manoeuvre distance used by Yoshimura and Soeda (1972) and Miller (1981).

Once the manoeuvre has been detected, a compensation is applied to the tracking filter to allow for the unexpected behaviour. The mildest action that can be taken is to increase the filter error covariance, and so increase the filter gain. Thus the filter becomes more sensitive to its input data and is enabled to lock onto the new target track. Yoshimura and Soeda adjust the filter covariance dependent on the innovation that triggered the manoeuvre test; Demetry and Titus reset the Kalman filter gain to an early point in its sequence; Holmes uses the same technique, modified for use with α-β filters. Demetry and Titus also reprocess the last few measurements using the adjusted gains.

A more drastic approach is to reinitialise the tracking filter, either from scratch or making use of the information included in the discarded track. This method is used by Miller (1981) in his multiple option tracker.

An original implementation of this idea was suggested by Pardini and Grasso (1973). They attached two tracking channels to each target, one with a large transition noise and the other with a small transition noise. The former channel has a short time constant, so it is able to follow sharp manoeuvres; the latter channel, with its high spatial resolution, provides more precise state estimates. When a manoeuvre is detected the high precision channel is re-initialised by copying from the low resolution channel. The method is also used in the Searchwater Radar system described by Symons (1982).

Finally, a novel manoeuvre handling approach was advanced by Bar-Shalom and Birmiwal (1982). Their variable dimension filter operates in a "quiescent" mode using a simple constant velocity state transition model, until a manoeuvre is detected. It then switches to a higher dimension state transition model that incorporates acceleration components, initialising the new state and error covariance estimates from the values obtained with the simpler model. The

filter reverts to its quiescent mode when the accelerations estimated by the tracking filter become statistically insignificant in relation to the estimation errors. Manoeuvre detection is handled by a fading memory average of the χ^2 manoeuvre distance.

The filter is computationally efficient and economical on storage since it uses a complex model of the target motion only when such a model is necessary, and it was found to function well in a computer simulation test reported by Bar-Shalom and Birmiwal.

These "detect and correct" methods provide good manoeuvre handling performance coupled with good tracking over non-manoeuvring sections of target trajectory at very little computational cost. They allow the system to use advanced track updating techniques on the non-manoeuvring parts of a trajectory, but to handle manoeuvres successfully as well. An illustration of this is the system described by Fortmann et al. (1983) which uses the JPDA track updating scheme.

7.5 Segmenting Sonar Input.

The algorithms reviewed in the preceding sections are all intended principally for use with radar or passive sonar sensors. In this section the differences between these environments and the context of active sonar interpretation will be explored and some suggestions will be presented for a segmentation scheme suited to sonar interpretation.

The environment in which the sonar interpreter must operate has all of the complications discussed above -- it is a multitarget, cluttered environment with uncertain detection and a proportion of manoeuvring targets. The continual movements of targets and observer will cause predictable and unpredictable obscuration; new targets will appear and old ones cease to be visible. Thus a segmentation scheme in this environment must have competence in all three areas reviewed -- it must be able to initiate and maintain tracks, and to compensate for or accommodate target manoeuvres. In common with radar, the signal to noise ratio and spatial resolution of active sonar can be made high, so that input data is cleaner than in the

passive sonar case; in common with passive sonar, the detection probability may be fairly low, so that missing data is relatively common compared with radar.

The most significant difference between the sonar interpretation context and the radar and passive sonar context is in the handling of stationary and slow-moving targets. In radar systems these targets, which arise mostly from ground clutter, are of little interest -- the system aims to track relatively fast-moving targets such as aircraft or missiles -- and are eliminated in a prefiltering stage. In passive sonar, a stationary target may have no sonar emission, since the emitted sounds are normally due to engine and transmission noises, and therefore be undetectable. Stationary targets are often of little interest in passive sonar applications.

In contrast, the sonar interpreter is intensely interested in stationary and slow-moving targets for two reasons: first, such targets are stable sources of echoes in the environment it is trying to model and are therefore significant in purely sensory terms; and second, it is on the stability of these sources that the system's sensory reference frame stability ultimately depends. The implication of this contrasting focus of attention is that the sonar interpreter system will tend to be more heavily loaded than an equivalent radar or passive sonar system, since it is interested in a larger proportion of its sensor traffic.

A second, related, difference concerns the trade-off between successful tracking of true targets and the generation of false target hypotheses. A segmentation system which liberally hypothesises new targets will be more successful at detecting and tracking true targets than a more conservative system. The cost of its success is that a (possibly large) number of false target hypotheses will be generated because of clutter detections.

How conservative a segmentation should be depends on the relative cost of false alarms and missed detections. In military or civil air traffic applications the cost of missing a true target is high; segmentation systems tend therefore to be liberal when postulating targets. For sonar interpretation, the converse is true -- a missing

target gives no information to the interpreter, whereas a false target hypothesis gives false information. The segmentation system will therefore be more conservative, and this will tend to reduce the system loading since fewer hypotheses will be entertained simultaneously.

A third major difference between the sonar interpreter context and the radar or passive sonar context is the availability of global information. In the latter context targets are, for tracking purposes, independent point objects. In the former context many echo sources (the "targets" that are being tracked) are expressions of the topology and geometry of their parent objects. Thus targets may not be independent, nor do they behave identically. For example, the data association properties of a linear source differ from those of a corner. (This parallels the model of track updating given by Alspach (1975) where there are several possible target types but returns are anonymous so the tracker has no certain knowledge of the type of target that generated a given set of returns.)

Higher level processing that attempts to identify the topology of echo sources can also provide useful information for tracking. The source topology could be used to select among transition models in the tracking filters or to choose between possible segmentation strategies. Reasoning about combinations of sources could be used to adjust segmentation parameters. For example, the presence of a pair of linear sources suggests a corner source at their intersection; this hypothesis motivates an increase in the new target density parameter in the neighbourhood of the intersection, where an as yet unobserved source is predicted.

7.5.1 A Modular Segmentation System.

In the context of a sonar interpreter, segmentation is the task of classifying input events with respect to the known echo sources recorded in the Target Database, postulating new sources as required. The postulated sources and accumulated history trees that describe the classification decisions taken by the segmentation process form

the interface between Viewpoint Registration and the later stages of the interpreter.

The segmentation process must be able to handle different classes of targets, for example, stable reference points for motion resolution, targets representing different source topologies, and targets whose underlying source topology is unknown or partially known. The system must also be able to vary the segmentation parameters, for example, the clutter and new target densities, in different regions of space, to model conditions known by the interpreter to exist in those regions. These requirements can be met by using a modular segmentation algorithm and a track status labelling scheme.

For general purpose track maintenance the most promising state-of-the-art algorithms are Reid's (1979) and the Joint Probabilistic Data Association method of Fortmann et al (1983). The clustering properties of these techniques provide the required modularity, with the possibility of different segmentation parameters for each cluster, and are well suited to exploit any hardware parallelism (each cluster could be implemented as a separate process).

Reid's work is particularly appropriate in our context for two reasons. First, his algorithm includes automatic track initiation within the same framework as track updating (whereas JPDA assumes that track initiation has been done); and second, the algorithm stresses the explanation of events, rather than the observation of targets, and employs an explicit history representation.

The major overhead with the clustering algorithms is the formation of the clusters, potentially $O(n^2)$ in the number of targets. This is aggravated in the sonar interpretation context by the higher density of interesting targets (the actual number of targets to be tracked could be adjusted by varying the sensitivity parameters in the Sonar device Interface, but a useful system would have to accommodate about 50 to 80 targets). The difficulty can be overcome, however, by incorporating a spatial indexing mechanism into the Target Database, so that the near neighbours of any target are accessible at once given a reference to the target.

The track status labelling scheme referred to above must therefore include the spatial indexing or clustering information for each target, as well as the target status (tentative or confirmed), the motion status for motion resolution (stationary, moving, manoeuvring, etc.), and the source topology information currently hypothesised by the Event Analysis process. The labels might also incorporate a reference to the appropriate segmentation parameters for the target or cluster.

There are two possibilities for handling unexpected manoeuvre. Reid's algorithm may be extended to consider target manoeuvre as a candidate explanation (possibly in the manner advocated in (Kenefic, 1981)), which leads to a more prolific growth of history with consequent demands on storage and processing resources. Alternatively, manoeuvres may be detected and compensated as they occur. This strategy is computationally less expensive and is easily accommodated in a modular processing structure. The methods of Bar-Shalom and Birmiwal (1983) or Pardini and Pardini and Grasso (1973) are especially attractive here, since they make good use of the existing data available for the manoeuvring target. Adaptive techniques such as that of Moose (1975) are less appropriate here because of the multiple tracking filters they employ.

The major constraint affecting the choice of manoeuvre processing strategy is that the motion resolution system state estimates must be perturbed as little as possible by the effects of manoeuvres. The former strategy is better than the latter in this respect, but the trade-off against the strong real-time constraint may favour the latter method. Ultimately the balance must be determined empirically.

Expected observer manoeuvre can be compensated by the motion resolving system and vehicle control system in such a way that its effects are not felt by the segmentation processing.

7.5.2 Summary.

The segmentation processing for a sonar interpreter can therefore be constructed as a composite of several of the state-of-the-art algorithms reviewed in this chapter. However, there are various possible realisations of the modular segmentation system, differing in power and computational cost. The final system adopted will depend on the practical complexity of the segmentation problem to be solved by a real interpreter (which depends on the proportion of clutter detections, the density of targets, and the proportion of manoeuvring targets) and on the computational resources available for the solution of segmentation. It remains a research issue for the future.

Chapter 8. Specular Event Analysis.

Specular events account for a significant proportion of all the echoes detected by a sonar device. The echoes arising from specular reflections are amongst the strongest available and in a world where fading is normal and noise is severe they are often the most stable and consistent. Specular events therefore constitute an important category of information for sonar interpretation. In this chapter the motivation and function of Specular Event Analysis are discussed; six types of specular event source are identified and their properties described; and a collection of techniques and suggestions for Specular Event Analysis are presented.

8.1 The Notion of an Event Source.

Before discussing the analysis of specular events it is necessary to formalise the notion of an event source. The term has been used informally to describe the entities in the environment that cause echoes and was used to denote the entities postulated by segmentation processing to explain tracks. The formal meaning chosen for the term is close to its intuitive meaning and covers both specular and diffuse reflections.

> An Event Source is a surface or configuration of surfaces, forming part of an environmental object, that reflects acoustic energy back to an insonifying observer, thereby causing event detections.

The events generated by a source are grouped together by the segmentation module into a track. The track can be considered to be a set of (separate) related events generated by the event source, or the trajectory of a moving point target. The former view is most useful for analysing the static properties of the event source (for example, its shape or position) while the latter is more appropriate when analysing the motion reflection properties of the source. The target motion is determined by the structure of the event source and

the proper motion of the observer. The distinction is ignored here, however, and the two conceptions of a track are used interchangeably.

Event sources consist of environmental surfaces in certain configurations. Thus there are two classes of property associated with a source: geometric properties and topological properties. The former comprise the specific sizes, shape parameters and locations of the surfaces that form the event source. The latter comprise the size and shape independent properties of the source -- those aspects of the source behaviour which depend on the qualitative interrelationships of the source surfaces and on the interaction of the source surface configuration with the viewpoint position.

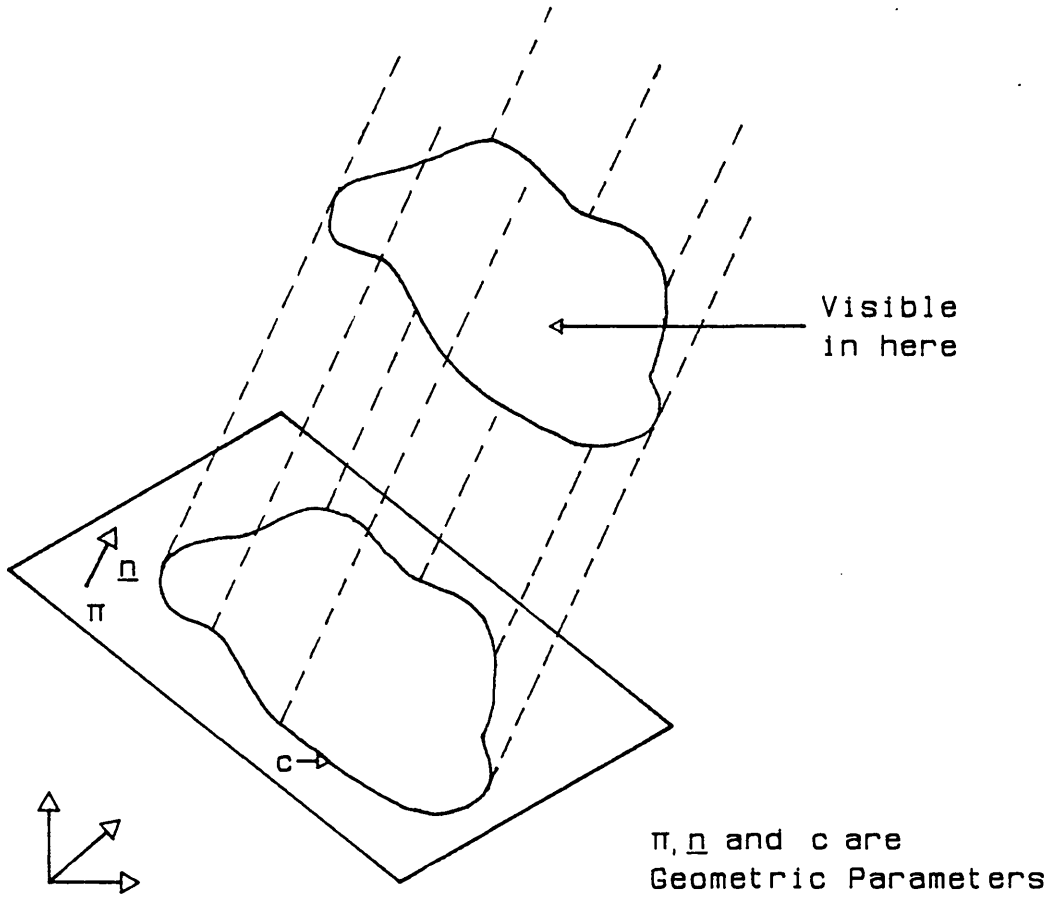


Figure 8.1. Geometric and Topological Source Properties.

An example illustrates and clarifies this distinction. In Figure 8.1 a source comprising a small planar surface patch is shown. The geometric properties of this source include the surface normal, $\underline{n}$, its location in space (fixed by a point on the plane, for example π), and a specification of the curve (c) bounding the surface patch. The

topological properties include the fact that the source is visible only at points within a cylinder of space and that the source surface is a convex set of points.

8.1.1 Event Sources as Interpreter Cues.

The task of Specular Event Analysis is to determine both the topological (source type) and geometric (surface parameter) properties of each source postulated by segmentation as the explanation of an event set. The motivation for this task is that the properties of sources are important cues for the interpretation process. When the properties of a particular source are known, the surface parameters and the disposition of surfaces in a neighbourhood on the underlying object are also known. This information is then available to drive the object modelling activity of the sonar interpretation system.

Knowledge of the source properties is useful in two other ways. First, the different source types reflect observer motion in differing ways depending on the source topology. The topological properties of the source may be used to construct a correction term for the motion resolution system to account for the source behaviour. Second, a partial knowledge of the properties of a given source may be used to suggest suitable observer motion strategies for elucidating the information necessary to complete the description of the source.

8.1.2 Related Work in Optical Flow.

Koenderink and van Doorn (1977) have considered the use of specular points as descriptive features in the problem of discovering permanent environmental structures from the optical flow field perceived by a monocular exploratory observer in a stationary environment. They explore the topic in the context of a theory of the environmental modelling activity performed by natural vision and perception systems.

The problem they address closely resembles the sonar interpretation problem. In both cases the exploratory observer must discriminate between proprio-specific sensory change induced by its own motion and extero-specific sensory changes corresponding to the effects of environmental objects. Extero-specific visual invariants are of particular importance since they relate to permanent structures in the environment.

The difference between the problems is threefold. First, the work of Koenderink and van Doorn assumes that a complete optical flow field is available as the starting point for its analysis, while sonar interpretation assumes a sparse apparent motion field. Second, the greater specularity of sonar imaging makes specular events more frequent and more easily identifiable than their optical counterparts. Third, the sonar interpretation problem admits environmental (extero-specific) motion as well as observer motion. Despite these differences the theory developed by Koenderink and van Doorn is directly relevant to sonar interpretation at the event analysis and environmental modelling levels.

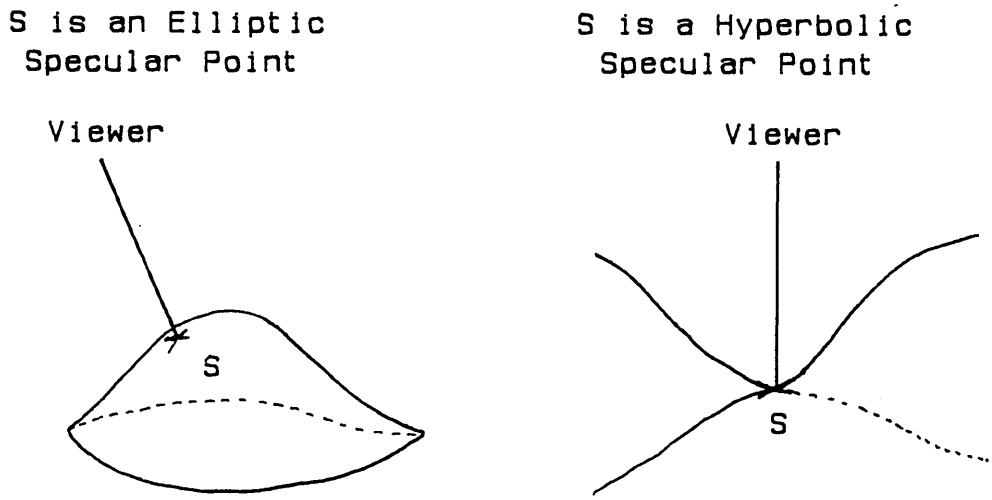


Figure 8.2. Examples of Specular Points.

Specular points are defined as those object surface points at which the local tangent plane is normal to the visual direction.[1]

1. By Koenderink and van Doorn, op. cit. page 235. This is an unusual definition in an optical context, where specular points are generally taken to be those points at which the observer's visual direction is aligned with the image of an illuminant. In our context

(Figure 8.2). The distance from observer to object surface at such a point is therefore either at an extremum (Figure 8.2a) or a point of inflexion (Figure 8.2b). The former, also called elliptic points, correspond to convexities or concavities while the latter, also called hyperbolic points, correspond to saddles in the object surface. It is known from the mathematical theory of global analysis that a smooth surface enclosing a volume (such as an object boundary surface) must possess at least one elliptic point but need have no hyperbolic ones.

Using global analytic techniques Koenderink and van Doorn show how an observer can construct a description of object shape in terms of the singularities of the object surface -- the specular points and occlusion boundaries. For any view of the surface, the observer computes the slant vector field (the gradient field of reciprocal observer to object point distance) from the optical flow, and a global analysis of this field provides a qualitative description of the surface as a graph of its visible singularities (nodes) and the field lines between them (arcs).

As the observer viewpoint changes, the singularities move relative to each other, and at certain distinguished viewpoints singularities are created or annihilated in pairs, altering the graphical description of the view. Differential topology teaches that the distinguished viewpoint set is of measure zero and partitions the space of all viewpoints into cells within which the global view description is stable under exploratory observer movements that do not cross a partition. The graphical description of the view is thus an exterospecific invariant inside the cell.

Koenderink and van Doorn go on to define the "visual potential" of an object as the transition graph of the invariant views. It is isomorphic to the cellular spatial partition, with a node for each stable viewpoint cell and an arc for each partition boundary. The visual potential allows an observer to predict the manner in which the view of an object will change in response to a planned movement

---

it is a natural definition, since the 'illuminant' is carried by the observer.

and to determine in which spatial cell the current viewpoint belongs.

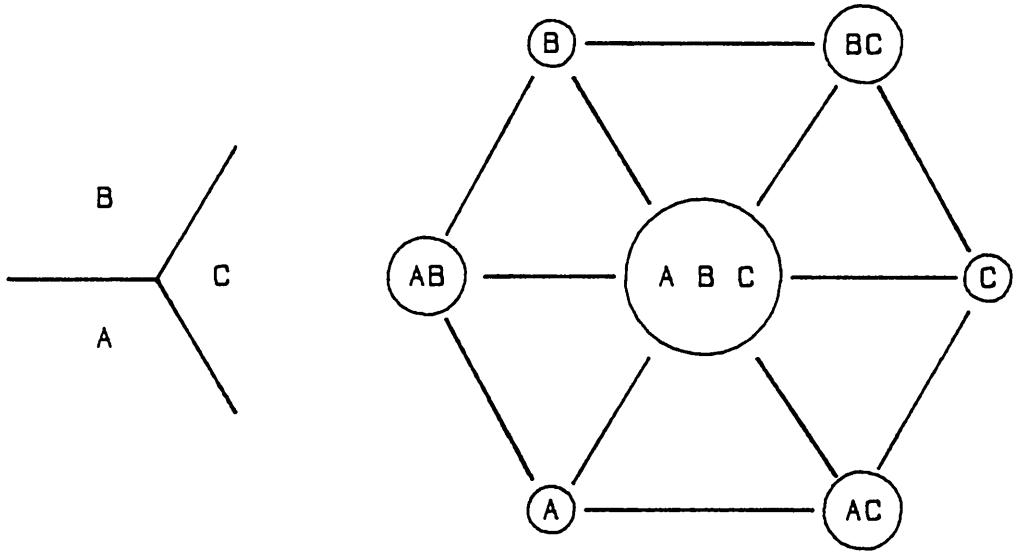


Figure 8.3. Visual Potential of a Convex Corner.

An example of a visual potential, to clarify the ideas of this section, is shown in Figure 8.3. It is a subgraph of the cube's visual potential exhibited in (Koenderink and van Doorn, 1977). (The graph of the visual potential shown as Figure 11 of that paper is incorrect: there are numerous edges missing, since the graph must be invariant under a renaming of the faces that preserves the topology of the cube.) The example chosen here is a convex corner. The corner point and the edges between faces are occlusion singularities and each face provides a specular point. The visual potential nodes describe what is visible in a particular stable view (in the figure they are labelled with the faces that are optically visible, to conform with Figure 11 mentioned above); the arcs define the allowed transitions between stable views. Comparison of Figure 8.3 with the corresponding figure in Chapter two, Figure 2.4, shows the isomorphism between the cellular partition of space and the visual potential graph.

8.2 Properties of Event Sources.

Event sources have three properties potentially useful for classification and recognition of source types. The imaging process involved in specular reflection constrains the visibility of a source, the positions of events it generates, and the proper motions of those events.

Source visibility is constrained by self-occlusion or by the alignment of observer and object surfaces required to permit a specular reflection. Thus for any source there is a subset of the space of all possible viewpoints such that the visibility constraints are satisfied whenever the observer position is within that set. This set of points, from which the source is theoretically visible, is termed the source visibility set.

The source visibility set concept corresponds closely to the visual potential of the source. For a single, isolated source the visual potential is a graph consisting of a pair of nodes connected by an arc. The nodes correspond to the visibility set and its set complement; the arc represents the boundary of the visibility set at which the catastrophic transition from visibility to invisibility occurs. Thus the visibility set contains all the information present in the visual potential graph for the source.

The source visual potential is not a particularly useful description of the source for discrimination purposes -- all sources have the same visual potential graph! However, the transition representation in the visibility set does indicate the source structure.

Events generated by a source are distributed in space in a manner that depends both on the nature of the source being observed and the position and orientation of the observer when each event occurs. The constraint in force here is that the positions of events generated by a source must lie on some surface associated with their generator. The set of event positions that a source may generate thus characterises the spatial configuration and size of the source; it will be termed the source surface set.

The third source property is a corollary of the event position constraint. Events may move only in ways that satisfy the position constraint of their generator: their trajectories must lie wholly within the surface set of the source responsible. This constraint applies locally, so that the instantaneous proper velocity of an event generated by a stationary source must lie tangentially to the surface set. If the surface set is a convex set it also applies globally, constraining the average velocity as well as the instantaneous velocity.

Another way of describing the motion constraint is to say that reflected observer motion depends both on source type and observer proper motion. This generalises the ideas of Chapters four and five, where all sources were treated as points. In general, unlike the point sources, a source will reflect observer motion only in certain specific directions determined by the source type and configuration. The events generated thus appear to possess a proper motion which precisely cancels the reflected observer motion expected by the motion resolution system in certain directions. This does not affect the correctness of the motion resolution systems, but it does affect their performance because events generated by stationary objects may have non-zero proper motions dependent on the generating structures, and some potential reference points are therefore moving.

8.3 Six Types of Specular Event Source.

In this section six classes of specular event source are defined and their properties -- their visibility and surface sets and their associated motion constraints -- are described. The sources are enumerated in order of echo strength per unit size. The enumeration therefore reflects the expected frequency of sighting for uniform environmental distribution of sources amongst the classes and for given source size.

8.3.1 The Concave Corner.

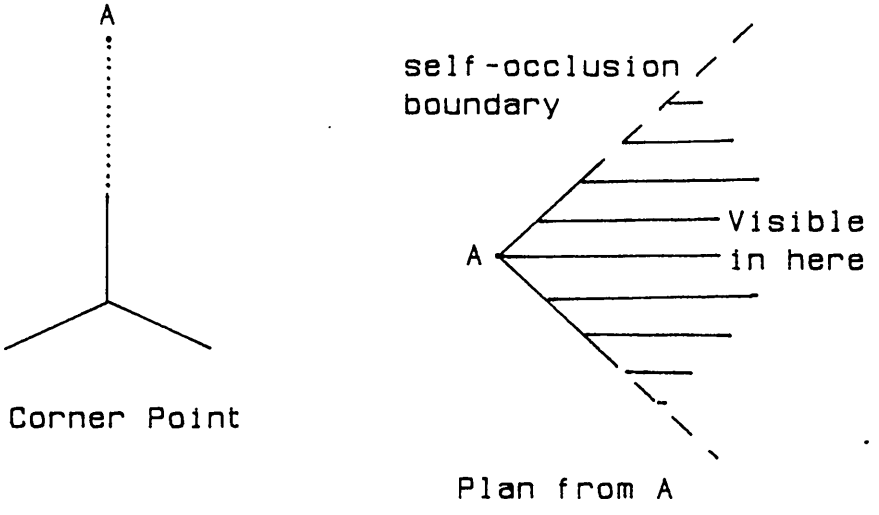


Figure 8.4. A Concave Corner Source.

By far the strongest echo source is the concave corner, where three locally planar surfaces meet at a point. This source is illustrated in Figure 8.4. Concave corner sources are common in the environment and tend frequently to be seen.

The visibility set of a concave corner is determined solely by self-occlusion. Neglecting diffraction effects the source is visible from any point to which a line-of-sight may be drawn from the intersection point of the corner surfaces (the corner point).

The surface set and associated motion constraint for a concave corner are particularly simple. The surface set is just the corner point, at which all events generated by the source appear to lie, and the source reflects all the observer motion.

The good visibility and stable position of corners make them important for the Viewpoint Registration sub-system. They correspond to the point sources used in Chapters four and five and would be the principal reference points used by the system in practice.

8.3.2 The Concave Linear Source.

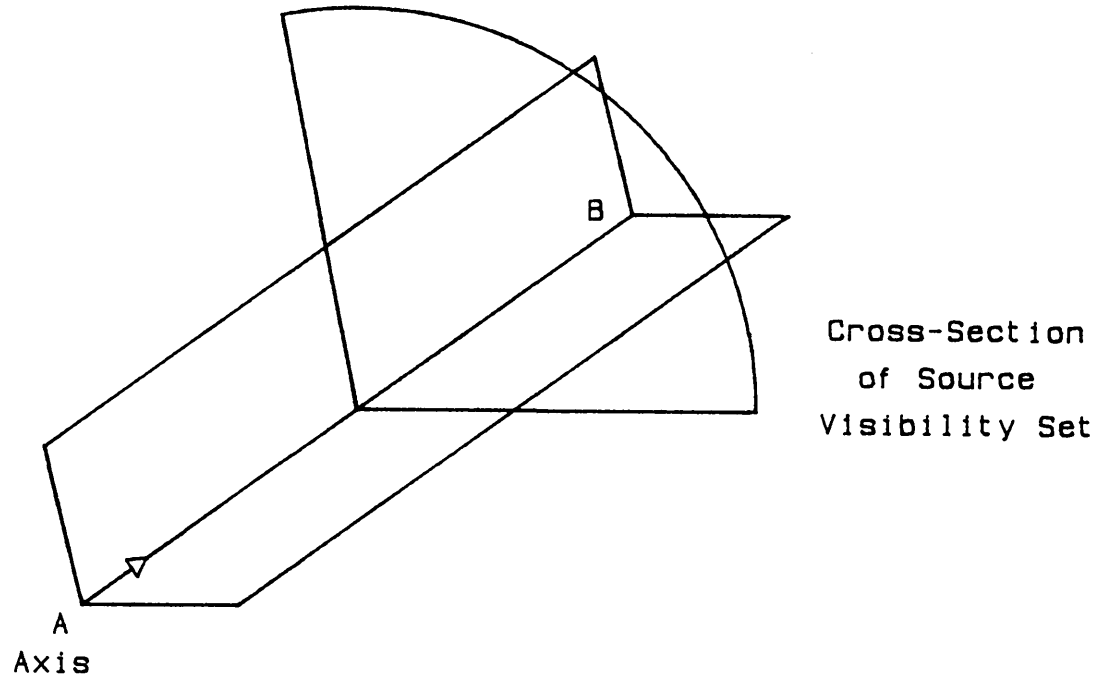


Figure 8.5. A Concave Linear Source.

The concave linear source is similar to the corner. It is formed where two locally planar surfaces meet along a line (the source axis) and is shown in Figure 8.5. The axis may be straight or curved (in which case the source is a curvilinear one).

Visibility for the concave linear source is determined by self-occlusion and by alignment constraints. The source is visible from all points from which a line of sight may be drawn between the (finite) source axis and the current viewpoint, meeting the axis at right angles. The visibility subset is thus the cylindrical sector of space occupied by the normal lines of sight.

The surface set for the concave linear source is the source axis. The source reflects observer motion orthogonal to the axis and absorbs observer motion parallel to the source axis direction. Events generated by this source therefore have an induced proper motion along the source axis equal to the axial component of the observer's linear proper motion. If the axis is not curved, the motion constraint applies globally as well as locally.

8.3.3 Planar Sources.

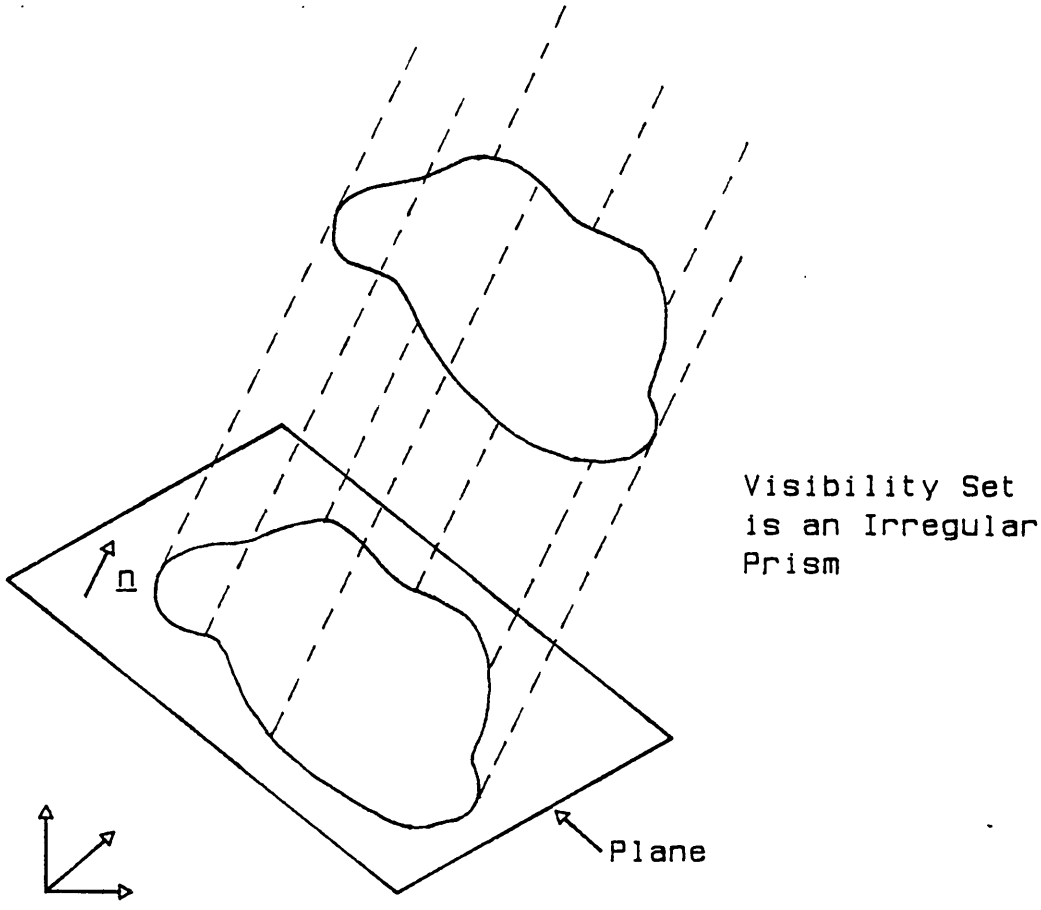


Figure 8.6. A Planar Surface Source.

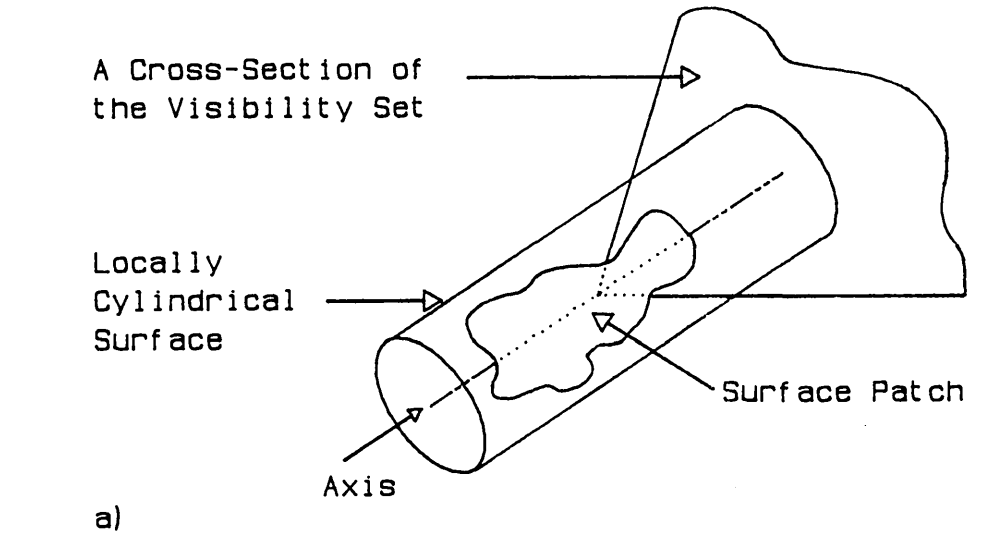
The third strongest source is the planar source (Figure 8.6) consisting of a planar surface patch.

The visibility set of this source is constrained entirely by alignment requirements. The source is visible whenever a line of sight from the viewpoint strikes the surface patch along the surface normal. The visibility set is thus the solid enclosed by the patch boundary curve swept along the outward surface normal.

The surface set in this case is the planar patch, and the source reflects observer motion only in the normal direction to the plane. Tangential motion is absorbed, so that events generated by this source have induced tangential proper motions equal to the projection of the observer proper motion onto the surface. The motion constraint applies globally for this source.

8.3.4 Cylindrical Sources.

These sources consist of a convex or concave surface patch on a circular or elliptic cylinder. Figure 8.7a shows a convex circular-cylindrical source.



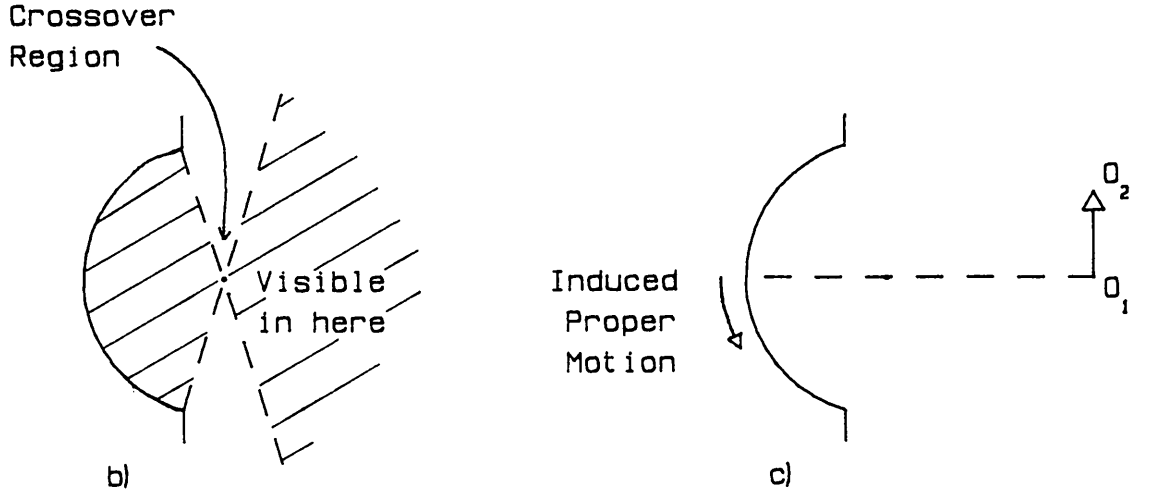


Figure 8.7. Cylindrical Sources.

The visibility set of the convex cylindrical source has a similar structure to that of the curvilinear source and contains all viewpoints from which a line of sight to the surface patch strikes the patch along its outward normal. For a circular-cylindrical source there is an associated curvilinear source with axis identical to the cylindrical source axis and bounding surfaces that match the

boundary of the cylindrical surface patch.

The visibility set of the concave cylindrical source is also defined as the solid comprising the space occupied by the normal lines of sight, but it has a characteristic structure, illustrated in Figure 8.7b in cross section for an elliptic cylindrical source. For a circular-cylindrical source the characteristic crossover region is confined to the source axis.

As for the planar source, the events generated lie on the surface patch of the source. However, the motion constraint is more complex. Observer motion normal to the surface is reflected, and axial motion is absorbed. Observer motion parallel to the curved surface causes an induced event proper motion with an angular velocity at the source axis equal to that of the observer's motion (for a concave source the induced motion is inverted with respect to the convex source). Figure 8.7c shows this induced motion for the elliptic-cylindrical concave source of Figure 8.7b. Since the surface is curved, the motion constraint does not apply globally to observer motion that contains any component parallel to the curved surface.

8.3.5 Spherical and Ellipsoidal Sources.

A third class of surface source is the class of spherical and ellipsoidal sources caused by surface patches with two finite radii of curvature. These sources may also be either convex or concave. A convex spherical source is illustrated in Figure 8.8.

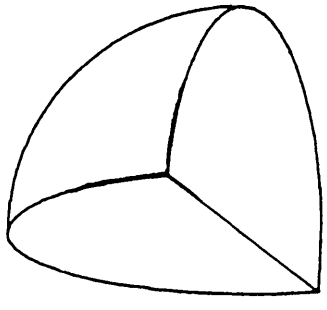


Figure 8.8. A Spherical Source.

The visibility set for a convex or concave spherical or ellipsoidal source is analogous to the corresponding cylindrical case; it comprises the space occupied by the normal lines of sight. For spherical sources there is an associated corner such that the bounding surfaces of the corner match the bounds on the spherical surface patch and the corner point is the centre of the sphere. The visibility sets for the concave sources have characteristic crossover regions analogous to those of the concave cylindrical sources.

The surface set for this source is again the surface patch itself and the motion constraint is analogous to that for cylindrical sources: observer motion normal to the surface is reflected, while motion parallel to the surface induces an event proper motion that depends on the observer range and the surface shape. For concave sources the induced proper motion is inverted.

8.3.6 Convex Linear and Corner Sources.

The sixth and weakest category of sources considered here are the convex sources analogous to the corner and curvilinear sources already mentioned. These arise because of weak reflection from object convexities (the reflections are generated by the mechanism described in (Freedman, 1962)) and are illustrated in Figures 8.9 (the corner) and 8.10 (the line source). In practical situations these sources are likely to be too weak to be distinguishable.

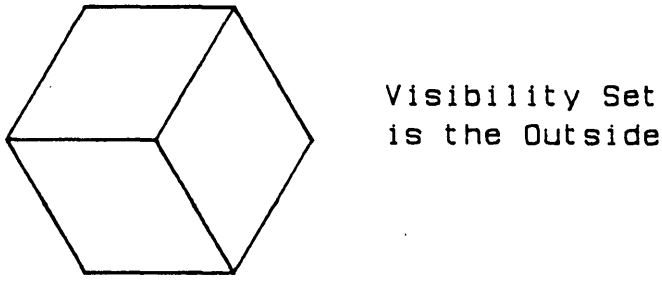


Figure 8.9. A Convex Corner Source.

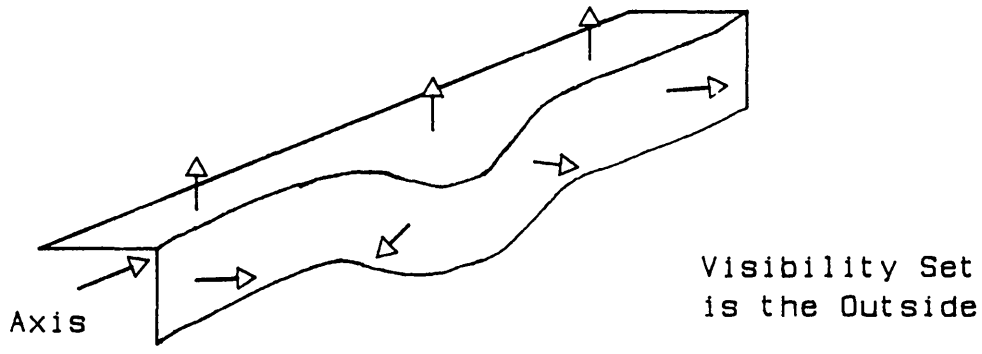


Figure 8.10. A Convex Linear Source.

The visibility set of the convex corner is determined solely by self-occlusion constraints and is the set complement of that of the corresponding concave corner. The visibility set for the convex (curvi)linear source is the region of space comprising all normal lines of sight from the source axis. Surface sets and motion constraints are identical with those of the corresponding concave sources.

8.3.7 Summary.

The properties of sources useful for recognition or discrimination are topological in the sense defined in section 8.1 -- that is, they depend on the general disposition of the source surfaces rather than the actual size or precise structure of each individual source. Each source type represents a family of sources of varying size and structure.

Source types are interrelated by their visibility and surface set properties. Thus planar, cylindrical, spherical and ellipsoidal sources group together as surface sources: their surface sets (which are two dimensional) are equal to the surface patches realising the sources and their visibility sets are defined solely by alignment constraints. The motion constraint for surface sources is that observer motion parallel to the surface set induces a similar (convex sources) or inverted (concave sources) proper motion in the generated events. This motion may partially or completely cancel the reflected observer motion (for the convex sources).

The convex and concave linear sources also group together, having a one-dimensional surface set. The corner sources, with a zero-dimensional surface set, form a third class.

Using the visibility properties, corner and convex spherical or ellipsoidal sources are grouped, as are convex cylindrical and linear; planar sources, concave cylindrical sources, and concave spherical or ellipsoidal sources each form a separate group. In the first case, the visibility set resembles a spherical sector of space, and corners may be thought of as small spherical sources. In the second case the visibility set resembles a cylindrical sector, and linear sources can be regarded as small cylindrical ones. For planar sources the visibility set is a prism, because of the determining role played by alignment constraints. Finally, the visibility sets in the last two groups are analogous to those in the first two, but contain crossover regions.

8.4 Source Recognition.

The Specular Event Analysis module of a sonar interpreter has two functions. First, it must discriminate among the various source types to identify the generating source type for each event. Secondly, it must extract from the events and motion associated with each source the source geometric properties (its corner point, axis or surface parameters) that determine a particular instance of the source family. This dual task is complicated by the inherent ambiguity of events, the proper motion of objects, and by fading and environmental noise.

Given complete and perfect information the topological properties of stationary sources are sufficient to discriminate each source type. Although the visibility and surface set properties are individually ambiguous, in combination as two independent dimensions of variability they uniquely indicate the source type. Table 8.1 lists the discrimination rules for stationary sources.

Visibility Set	Surface Set: Point	Surface Set: (Curvi)Linear	Surface
Prism	-	-	Planar
Cylindrical Sector	-	Convex or Concave (Curvi)linear	Convex Cylindrical
Cylindrical Sector + Crossover	-	-	Concave Cylindrical
Spherical Sector	Convex or Concave Corner, or Non-specular	-	Convex Spherical or Ellipsoidal
Spherical Sector + Crossover	-	-	Concave Spherical or Ellipsoidal

Table 8.1. Stationary Source Discrimination Rules.

There remains ambiguity between corner sources and non-specular (diffuse) echo sources. This is to be expected, since diffuse events are generated by object surface irregularities which are really small corner sources.

In practice, however, the information available to the system is neither complete nor perfect. A source will not in general be visible from every point in its visibility set because of random intermittent fading or occlusion by other objects in the environment. Thus only a subset of the visibility set, perhaps only a sparse subset, will generally be available. Motion of the objects on which the sources lie will affect the visibility and surface sets obtained for that source because the actual visibility and surface sets are moving in space. Measurement noise also smears the surface sets, turning corner sets into small ellipsoids, axial sets into small cylinders, and surfaces into thin laminae.

In view of this, source recognition is a difficult problem. No solution to the problem is presented here; rather, suggestions for the construction of a solution in terms of reasoning about the source surface and visibility sets are given. The characterisation of the

surface and visibility sets of sources is discussed, and an analytic technique for determining the topology of the surface set and for estimating the geometric parameters of a source is given.

8.4.1 Specifying the Source Reasoning System.

Partial analyses obtained using analytic or heuristic tests, applied to the available event sample for a source, motivate and support particular hypotheses about source type. Source discrimination within this context may then be handled by a source reasoning system able to combine the evidence from the various analyses that militates for and against each conjecture about a given source and able to compute a confidence assessment of the truth or falsity of each hypothesis. A suitable specification for such a source reasoning system is presented below. The system has not been implemented and remains a topic for future research.

System Goals.

The source reasoning system shares the two goals of Specular Event Analysis. It must determine the type of each postulated echo source recorded in the Target Database by the Viewpoint Registration system and it must extract geometric parameters appropriate to that source type. Each conjecture of source type is stored as an individual hypothesis with attached confidence information. The system may entertain multiple hypotheses about a given source at any time.

System Input.

Source reasoning is based on the outputs of the Viewpoint Registration module. Its input comprises two sets of data: (i) the contents of the Target Database, which provide descriptions of the track status of each source postulated by segmentation processing; (ii) the recorded tracking histories of sources, which provide positions and local motions of events associated by segmentation with each source.

Target Database input accumulates as tentative sources are confirmed. History input grows as new events are explained by already postulated sources.

Note that source reasoning is not a prerequisite for segmentation, although both processes are concerned with hypotheses about events and event sources. The segmentation processing determines that sources must exist in order to explain the presence of events; source reasoning assumes the presence of sources and tries to recognise them.

System Operation.

Source reasoning must function incrementally. It must verify existing hypotheses as new data becomes available. It must incrementally recompute the confidence limits on hypothesis validity and the associated geometric properties of each hypothesis. New hypotheses must be conjectured when new sources are postulated and the system must be able to handle the large initial uncertainties entailed by these. Sufficiently improbable hypotheses must be discarded.

The desired operation can be achieved by using a rule-based inference system embodying the relationships set out in Table 8.1. A strategy of conjecture and verify or conjecture and refute may be used as appropriate. The evidence combination functions can be handled using a probability-based confidence scheme. For this application, however, Bayesian theory is inappropriate: it is difficult to describe absence of knowledge in a satisfactory manner using the Bayesian theory, causing problems in dealing with new sources. Dempster-Schafer theory (Schafer, 1976) overcomes this difficulty and affords other advantages: it is possible to iterate over the hypothesis-evidence network to obtain better confidence estimates; evidence combination is commutative so the order of reasoning is not critical; and the theory is well suited to incremental computation if efficiently implemented. (The use of Dempster-Schafer theory for evidential reasoning is described by Wesley (1983), for example.)

8.4.2 Characterising the Source Surface Set.

The analytic technique to be described is a statistical method for characterising the dimensionality of the source's surface set. The result of this partial analysis provides one component of the data required to drive the discrimination rules of Table 8.1.

The method is incremental and operates by attempting to fit an optimal plane (using a least square error fit) to the events currently associated with the source. A side-effect of the least squares computation is an indication of the dimensionality of the convex hull of the sample of surface set, and thus of the dimensionality of the full source surface set.

Convex Hull Dimensionality	Interpretation
Zero	Corner or Non-Specular Source.
One	Straight Line Source.
Two	Curvilinear or Planar Source.
Three	Cylindrical, Spherical or Ellipsoidal Source.

Table 8.2. Significance of Sample Set Dimensionality.

For perfect data the sample set dimensionality can take any value from zero to three. Assuming that the set of viewpoints at which the sample members were obtained does not comprise only coplanar points, the implications of each value are tabulated in Table 8.2. Position noise in the events will in general spread the sample set out with respect to the true surface set and the dimensionality will always be three. It is still possible, however, to estimate the dimensionality of the surface set from the sample. Clearly the analysis is only possible if at least four events are available: the sample set dimensionality is otherwise at most one less than the number of events.

To fit an optimal plane to the sample set, let $\underline{l}$ and $\underline{m}$ be unit orthogonal vectors lying in the plane and let $\underline{p}$ be any point on the plane. If the sample set comprises points $\underline{p}_i$, i=1..N then, referring

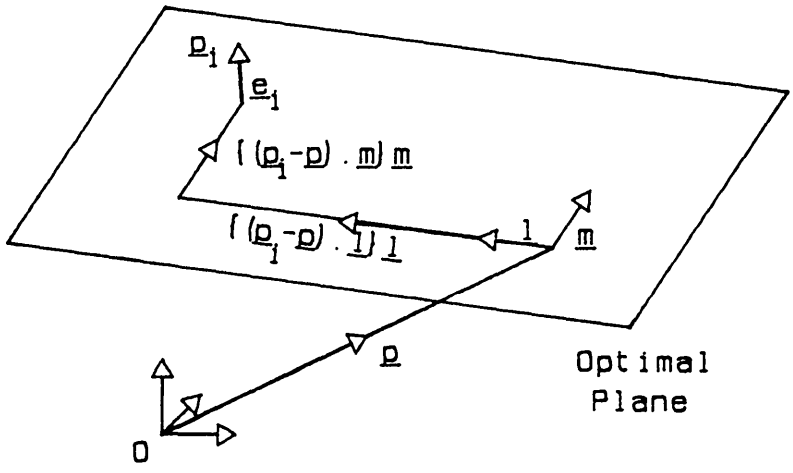


Figure 8.11. Fitting an Optimal Plane.

to Figure 8.11, the error residuals, which are vectors normal to the plane and passing through a $\underline{p}_i$, are given by

$$\underline{e}_i = (\underline{p}_i-\underline{p}) - \{(\underline{p}_i-\underline{p}).\underline{l}\}\underline{l} - \{(\underline{p}_i-\underline{p}).\underline{m}\}\underline{m}, \qquad (8.1)$$

and the total square error ρ is

$$\rho = \sum_{i=1}^{i=n} \left|\underline{e}_i\right|^2 = \sum_{i=1}^{i=n} \{ \left|(\underline{p}_i-\underline{p})\right|^2 - \{(\underline{p}_i-\underline{p}).\underline{l}\}^2 - \{(\underline{p}_i-\underline{p}).\underline{m}\}^2 \}. \qquad (8.2)$$

A standard analysis using Lagrange multipliers to obtain values of $\underline{l}$, $\underline{m}$ and $\underline{p}$ for the best-fit plane gives

$$\underline{p} = \frac{1}{n}\sum_{i=1}^{i=n} \underline{p}_i \qquad \text{(the centroid)} \qquad (8.3)$$

and $\underline{l}$ and $\underline{m}$ are eigenvectors of the matrix

$$S = \frac{1}{n-1}\sum_{i=1}^{i=n} (\underline{p}_i-\underline{p})(\underline{p}_i-\underline{p})^T = \frac{1}{n-1}\sum_{i=1}^{i=n} \underline{p}_i\underline{p}_i^T - \frac{n}{n-1}\,\underline{p}\,\underline{p}^T. \qquad (8.4)$$

The matrix S is the central scatter. It is the matrix analogue of the variance of a sample.

The dimensionality of the sample set is equal to the number of non-zero eigenvalues of the central scatter matrix. If all the members of the sample are coincident the central scatter is zero and the centroid is also coincident with the sample members. For a collinear sample set the eigenvector with a positive eigenvalue is the

line direction vector and the centroid is a point on the line; for coplanar events the eigenvector with zero eigenvalue is the plane normal direction.

The central scatter matrix can be computed incrementally from the history of events associated with a source by storing the number of events, the vector sum of their positions, and the sum of the position outer products $\underline{p}_i\underline{p}_i^T$. Whenever the scatter matrix is required it can be computed using the second form of equation (8.4).

When position noise is present, the centroid and scatter computed from a sample set will differ from the true centroid and scatter of the unperturbed events. The zero eigenvalues of the scatter matrix will instead be relatively "small" and the non-zero ones relatively "large". To quantify these vague relative expressions, consider the expected value of the central scatter $\overline{S} = \mathbf{E}\left[S \right]$.

Suppose that each point in the sample set is corrupted by an independent noise vector $\underline{n}_i$ satisfying

$$\mathbf{E}\left[\underline{n}_i \right] = \underline{\nu} \qquad \mathbf{E}\left[\underline{n}_i\underline{n}_i^T \right] = \mathbf{cov}\left[\underline{n}_i \right] + \underline{\nu}\,\underline{\nu}^T = \mathbf{N}. \tag{8.5}$$

The noise vectors are mutually independent. Then, using the definition of S in equation (8.4), the expected central scatter is

$$\begin{aligned}\overline{S} &= \frac{1}{n-1}\,\mathbf{E}\left[\sum_{i=1}^{i=n} (\underline{p}_i+\underline{n}_i)(\underline{p}_i+\underline{n}_i)^T - \frac{1}{n}\sum_{i=1}^{i=n}\sum_{j=1}^{j=n} (\underline{p}_i+\underline{n}_i)(\underline{p}_j+\underline{n}_j)^T \right] \\ &= \frac{1}{n-1}\sum_{i=1}^{i=n} \underline{p}_i\underline{p}_i^T - \frac{1}{n(n-1)}\sum_{i=1}^{i=n}\sum_{j=1}^{j=n} \underline{p}_i\underline{p}_j^T + \mathbf{N} \\ &= S + \mathbf{N} \end{aligned} \tag{8.6}$$

after a little algebra.

The effect of position noise is thus to add a term equal to the position noise covariance matrix to the true central scatter, and similarly the expected sample centroid $\mathbf{E}\left[\underline{p} \right]$ is displaced by $\underline{\nu}$:

$$\overline{S} = S_{true} + \mathbf{N} \qquad \mathbf{E}\left[\underline{p} \right] = \underline{p}_{true} + \underline{\nu}. \tag{8.7}$$

Estimates of $\underline{\nu}$ and **N** are available from the Viewpoint Registration sub-system -- they are the parameters of the absolute state error for the target associated with the events. In general, $\underline{\nu}$ is expected to be close to zero and **N** is just the estimated error covariance of the target absolute state. The true central scatter is a positive semi-definite matrix and **N** is positive definite, so the eigenvalues of the computed scatter are bounded below by the smallest eigenvalue of **N**.

The size of a "large" eigenvalue is illustrated below by means of a pair of examples given below: a linear source with axis passing through $\underline{p}_s$ and parallel to the unit vector $\underline{l}$; and a coaxial cylindrical source of radius a.

Line Source.

Each event in the sample set has a true position expressible as $\underline{p}_s + \lambda_i \underline{l}$ (where $\underline{l}$ is a unit vector) for some value of the scalar parameter λ_i. λ_i takes values dependent on the position of $\underline{p}_s$ and the length of the source axis.

The true central scatter in this case is given by:

$$S_{true} = \mathbf{var}\left[\lambda_i \right] \underline{l}\,\underline{l}^T. \qquad (8.8)$$

The relationship of noise covariance and true central scatter in the expected computed scatter matrix is thus independent of the size of the sample set (there is no explicit dependence on n, the sample set size). The eigenvalue of the true scatter, which is a lower bound on that of the computed scatter, depends upon the distribution of viewpoints from which the sample set was obtained and on the square of the length of the source axis.

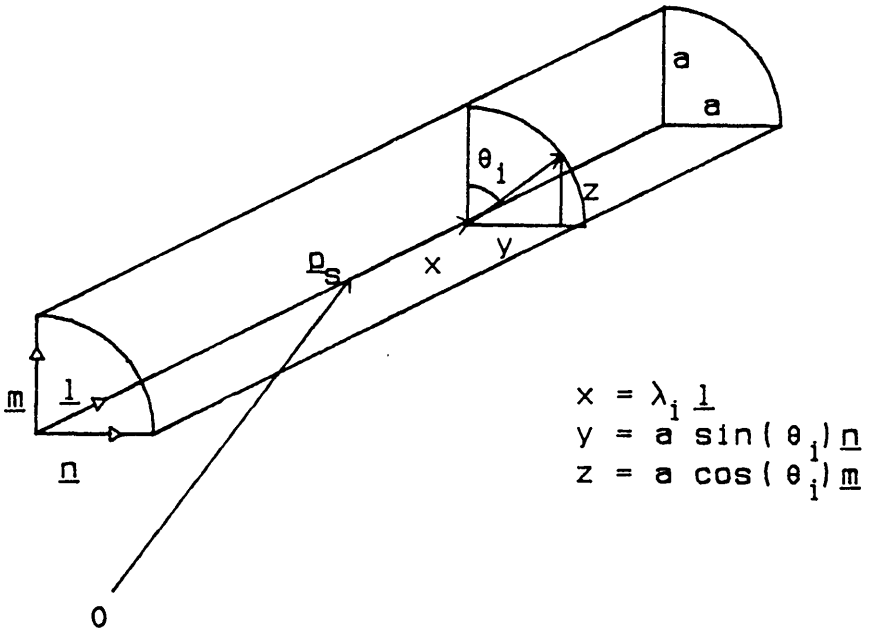


Figure 8.12. Parameterisation of Cylindrical Sample Set.

Cylindrical Source.

Let $\underline{m}$ and $\underline{n}$ be orthonormal vectors orthogonal to $\underline{l}$. The true position of the i^{th} event in the sample set is expressible as (see Figure 8.12)

$$\underline{p}_i = \underline{p}_s + \lambda_i \underline{l} + a\cos(\theta_i)\underline{m} + a\sin(\theta_i)\underline{n}$$

for some values of the scalar parameters λ_i and θ_i. The true scatter for this source is given by :

$$\begin{aligned} S_{true} = \quad & \mathbf{var}\left[\ \lambda_i\ \right] \underline{l}\,\underline{l}^T \\ & + a^2\mathbf{cov}\left[\ \lambda_i\cos(\theta_i)\ \right](\underline{l}\,\underline{m}^T + \underline{m}\,\underline{l}^T) \\ & + a^2\mathbf{cov}\left[\ \lambda_i\sin(\theta_i)\ \right](\underline{l}\,\underline{n}^T + \underline{n}\,\underline{l}^T) \\ & + a^2\mathbf{cov}\left[\ \cos(\theta_i)\sin(\theta_i)\ \right](\underline{m}\,\underline{n}^T + \underline{n}\,\underline{m}^T) \\ & + a^2\mathbf{var}\left[\ \cos(\theta_i)\ \right] \underline{m}\,\underline{m}^T \\ & + a^2\mathbf{var}\left[\ \sin(\theta_i)\ \right] \underline{n}\,\underline{n}^T. \end{aligned} \qquad (8.9)$$

As before the relationship between noise covariance and the true central scatter is independent of sample set

size. The true scatter matrix has three non-zero eigenvalues, one proportional to the square of the axis length, and two proportional to the square of the source radius.

These examples suggest that the eigenvalues of the central scatter matrix associated with a source will depend on the square of the source size. Thus the smallest detectable size of a feature will depend on the noise standard deviation, which is available from the noise covariance matrix **N**. A suitable strategy for analysing the central scatter is therefore to compare the eigenvalues of the computed central scatter matrix for each source with the estimate of the position error covariance available from the Viewpoint Registration system. Small eigenvalues should be of comparable size to those of the error covariance estimate, and large ones should be significantly bigger.

In practice, the estimate of **N** obtained from Viewpoint Registration processing will be different from the true **N** for two reasons. First, the upper bounds on error covariances used to simplify the combination of estimates will cause the estimated value of **N** to exceed the true value. Second, the estimate of **N** is mainly determined by the current relative position of the event being tracked (since the error covariance is ultimately calculated from a measurement error covariance whose directional dependence is a function of the relative position of the event), whereas the noise in the central scatter involves contributions from all along the track.

The usefulness of the estimate of **N** obtained from Viewpoint Registration is investigated briefly in section 8.5. An alternative possibility of using the central scatter at corner sources as an estimate of **N** is also considered there.

8.4.3 Determining the Structure of the Visibility Set.

The characterisation of the source visibility set is an inherently harder problem than determining the structure of the surface set and is not amenable to the same style of statistical technique. In this section the nature of the visibility set characterisation problem is discussed and some suggestions for tackling the task are made.

The decision implied by visibility set characterisation is a threefold choice. The region of space under consideration may be a prism, corresponding to a planar source; it may be a cylindrical sector, indicating a linear or cylindrical source; it may be an ellipsoidal sector, suggesting a corner or ellipsoidal source. Crossover regions in the visibility sets must also be detected. The basis for the decision is a sample set of viewpoints associated with events in the surface sample set.

There are two difficulties inherent in the problem. First, the sample set must be an adequate representative of the true visibility set if the decision is to be validly based. The sample set must not be artificially constrained -- for example, all the viewpoints coplanar -- with respect to the visibility set in question. It must also sample a substantial volume of the true visibility set. This latter requirement constitutes a serious problem for the sector visibility sets, where the volume of the set increases with radial distance squared or cubed.

The second difficulty is this: it is not possible to observe the "invisibility set" of a source. Because of the tendency to fading inherent in sonar, the fact that a source is not seen from a particular viewpoint does not imply that it is actually invisible there (although a viewpoint from which the source is never seen is very likely to be outside the visibility set). This feature of the problem makes it difficult to determine the boundaries of the visibility set.

A corollary to this point is that reasoning based on a conjecture and refutation strategy must of necessity postulate a restrictive visibility set for refutation, since only sightings of sources constitute certain refutation evidence. Viewpoints at which the source is invisible contribute evidence whose weight depends on the fading statistics of the source -- these are kept by the segmentation module in order to calculate the probabilities of event explanations.

The most appropriate strategy for the problem appears to be the construction, by the source reasoning system, of quantitative hypotheses describing the extent of space occupied by the visibility sample set under the assumption of particular source types. The possible hypotheses for a given source are constrained, once the surface set characterisation is known, by the rules of Table 8.1. The viability of hypotheses is assessable using the certain evidence of new source sightings and the uncertain negative evidence of non-detection. The distance from observer to source is also relevant in this reasoning process since it determines the scale of the visibility sample set.

Further constraints on the set of viable hypotheses for a given source are obtainable from a consideration of the vectors from sample viewpoints to their corresponding event positions. If all such vectors are roughly parallel, the visibility set is probably a prism; if the vectors all converge then the visibility set is probably a spherical sector; if they comprise skew lines converging into a linear region then the visibility set is probably a cylindrical sector.

Finally, the presence of a crossover region in the visibility set may be deduced from the inverted induced motion of the events as the observer moves or from the consideration of the view vectors (if the intersection region is significantly in front of the source it is probably a crossover region).

The outputs of such a strategy are twofold. First, a tentative source visibility set classification is computed as part of the description of the sample set. Second, the reasoning system can make conjectures about the visibility set that are directly verifiable using an appropriate observer motion, and thus suggest exploratory

strategy for the vehicle path planning component in an intelligent control system.

8.5 Some Experiments with Surface Set Analysis.

Two experiments using the central scatter technique for surface set characterisation were done using the two-dimensional motion resolution system described in Chapter four. These experiments form a small feasibility study for the three-dimensional use of the method. In particular, they illustrate the relationship of the computed central scatter to the true value, and the accuracy of the estimates of source geometric parameters. They also indicate the usefulness of the estimated position error matrix in assessing the dimensionality of the sample set implied by the computed scatter matrix.

The only simplification in passing from three to two dimensions is a reduction in the variety of source types. In two dimensions there are essentially two -- the corner, either concave or convex, and the curvilinear source of which the straight line segment is a special

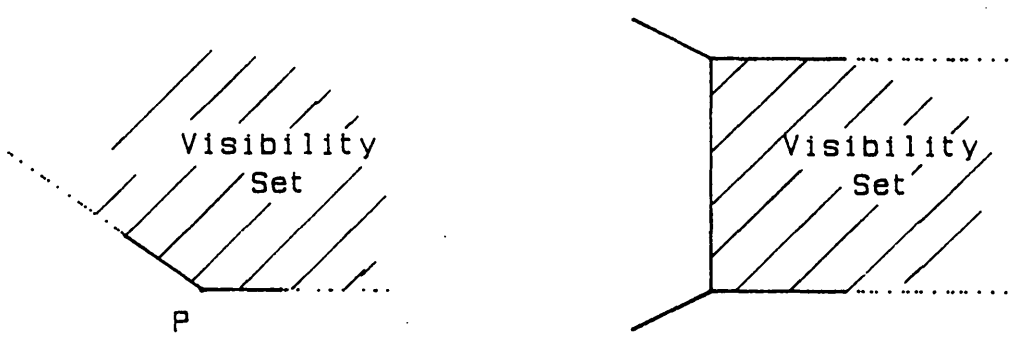


Figure 8.13. Two-Dimensional Source Types.

case. These sources are illustrated in Figure 8.13. The associated geometric parameters are the corner point for the corner source and the line parameters or equation for the curvilinear one.

The tests deal with two instances of two-dimensional sources. The source for the first test is a corner source at the origin. In the second test, a straight line segment is used. In both tests the observer moves in a linear path. The only noise present is

measurement noise, set at 10 milliradians standard deviation in the angular dimension and 1 centimetre standard deviation radially (chosen, as in section 4.2.2, to represent a medium resolution sonar sensor).

8.5.1 Test One -- A Corner Source.

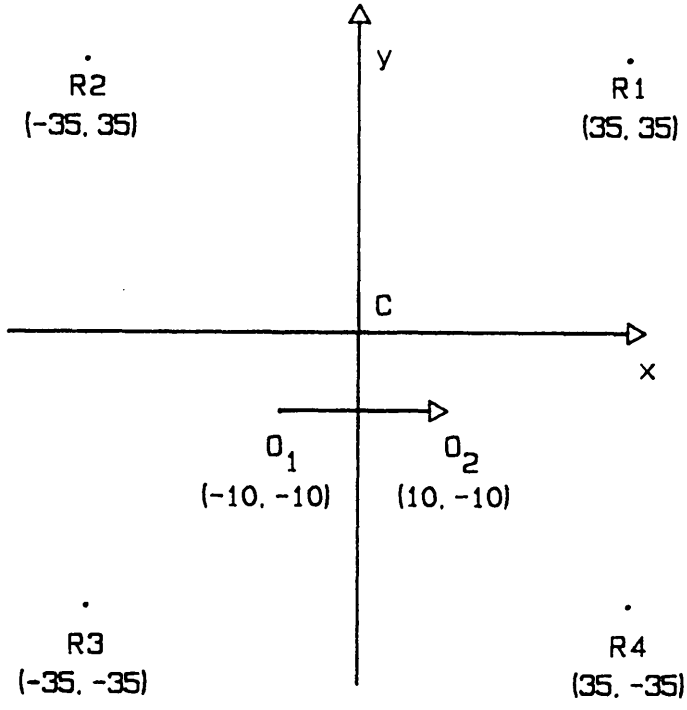


Figure 8.14. Test One Layout.

The layout of test one is shown in Figure 8.14. The observer moves from position O_1 at $(-10, -10)$ to O_2 at $(+10, -10)$ while observing the corner source at the origin. The whole trajectory $O_1 \rightarrow O_2$ is in the source visibility set, which is not otherwise specified. Four other targets arranged in a 70 metre square centred on the origin provide stable reference points for the viewpoint independent reference frame. The test runs for 60 seconds.

For the corner source the theoretical value of the true central scatter is zero, so the expected scatter is just the position noise covariance matrix.

Results.

The values of the computed scatter and of the estimated noise covariance are given below. The computed central scatter is based on 150 observations of the source.

$$\underline{p} = \begin{bmatrix} 0.010 \\ 0.007 \end{bmatrix}, \quad S = 10^{-4}\begin{bmatrix} 19.391 & 0.506 \\ 0.506 & 6.678 \end{bmatrix}, \quad \hat{P} = 10^{-2}\begin{bmatrix} 3.95 & 1.99 \\ 1.99 & 2.98 \end{bmatrix}.$$

The centroid computed during this test agrees to better than 1 cm with the true value, and the computed central scatter is also close to the true value (zero). Note that the central scatter computed in the test is direction dependent, with about three times more noise in the direction of observer motion than perpendicular to that direction.

The covariance estimated by the motion resolution system is much larger (by a factor of about 30) than the computed central scatter and, as expected, has a different directional dependence. The eigenvalues of the central scatter are 6.668 cm^2 and 19.401 cm^2.

Conclusions.

The test gives an indication of the size of the central scatter matrix to be expected for a corner source. It also suggests that the estimated error covariance matrix is not a useful standard of comparison for assessing the eigenvalues of the central scatter.

8.5.2 Test Two -- A Straight Line Source.

The source used in test two was a straight line source comprising a line segment passing through the origin. All other details were identical to those in test one. Figure 8.15 illustrates the arrangement of the observer and event source in this test. As the observer moves from O_1 to O_2 the event generated by the source travels uniformly from point L_1, at (-6, 2) to point L_2, at (12, -4).

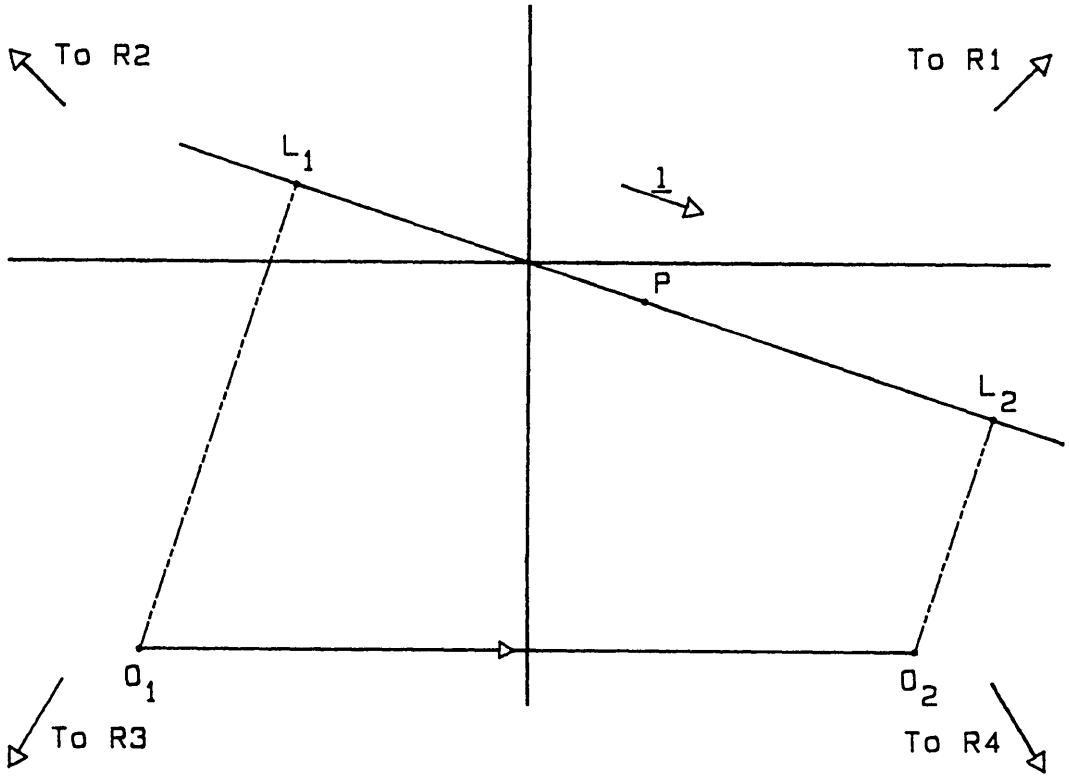


Figure 8.15. Test Two Layout.

For this source the central scatter given by theory is computed from a two-dimensional version of equation (8.8). The unit vector along the source axis is

$$\underline{l} = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 \\ -1 \end{bmatrix} \quad \text{and} \quad \underline{l}\,\underline{l}^T = \begin{bmatrix} 0.9 & -0.3 \\ -0.3 & 0.1 \end{bmatrix}.$$

The true centroid of the sample set for this source is at the point P, at (3, -1). The central scatter matrix is proportional to the variance of the parameter λ_i which in this case is uniformly distributed and, taking the centroid as $\lambda_i=0$, varies from $-3\sqrt{10}$ to $+3\sqrt{10}$. The scaling factor is thus 30 and the true central scatter is

$$S_{true} = \begin{bmatrix} 27 & -9 \\ -9 & 3 \end{bmatrix}.$$

Results.

The computed scatter and estimated position noise matrix obtained in this test are given below. As in test one, 150 events contributed to the computed scatter matrix.

$$\underline{p} = \begin{bmatrix} 2.948 \\ -0.928 \end{bmatrix}, \quad S = \begin{bmatrix} 26.776 & -8.884 \\ -8.884 & 2.949 \end{bmatrix}$$

As in test one, the computed and true central scatters matrices are in close agreement, and the true and computed centroids also agree closely. The eigenvalues of the computed scatter matrix are 29.73 m^2 (close to the true value of 30 m^2) and 0.0011 m^2 (for the true eigenvalue of zero), showing clearly that the surface set is one-dimensional. The eigenvector of the computed central scatter with eigenvalue 29.73 m^2 is $[0.9490\ -0.3151]^T$. The corresponding true eigenvector is $[0.9487\ -0.3162]^T$, again agreeing well with the computed value.

Conclusions.

In this case the computed central scatter and centroid give accurate estimates of the geometric parameters of the line source and the central scatter is strongly directional, correctly indicating a one-dimensional source surface set. The ratio of eigenvalues (more than 10,000) shows the extent of the difference between "large" and "small" eigenvalues. The good agreement in size of the small eigenvalue with the eigenvalues computed in test one support the idea of using the central scatter of corner sources as a standard for assessing the eigensolutions of central scatter matrices for other sources.

8.5.3 Summary.

The computation of central scatter as a method for analysing the properties of the source surface set seems promising from this small feasibility study. In each of the cases tested the computed centroid and central scatter were close to the expected values for the sources concerned and the difference between large and small eigenvalues was obvious.

The tests suggest that the central scatter computed for corner sources is a better standard than the error covariance estimated by the motion resolution system for assessing the size of the eigenvalues for other sources.

8.6 Conclusion.

The usefulness of the techniques discussed above for determining the nature of event sources depends on two factors: the specularity of sonar in practice; and the relative size of the environmental features that give rise to echoes, in comparison to the noise present in the estimates of event absolute positions. Objects in the marine environment are large enough to make possible some analysis of the sort suggested in this chapter, and real sonar data shows evidence of the type of specularity described. However, the topic remains an exciting possibility for future empirical research using real sonar data.

Appendix A. A Simple Kalman Tracking Filter Application.

This appendix illustrates the design and operation of the Kalman filter by describing a simple application -- the task of tracking the descent of a falling body. It is intended to show in a clear way the definition and function of the various mathematical components of the filter. A good introductory treatment of Kalman filtering is given by Mattin (1982) or Bozic (1979).

Figure A.1 depicts the tracking task chosen for this introduction. A package is dropped from a helicopter and is tracked by a ground station radar that can provide a regular series of noisy measurements of the height of the package above the ground. The tracking filter has the task of estimating the height and downward velocity of the package (for example, to predict when the package will hit the ground). The effect of wind and air resistance will be neglected.

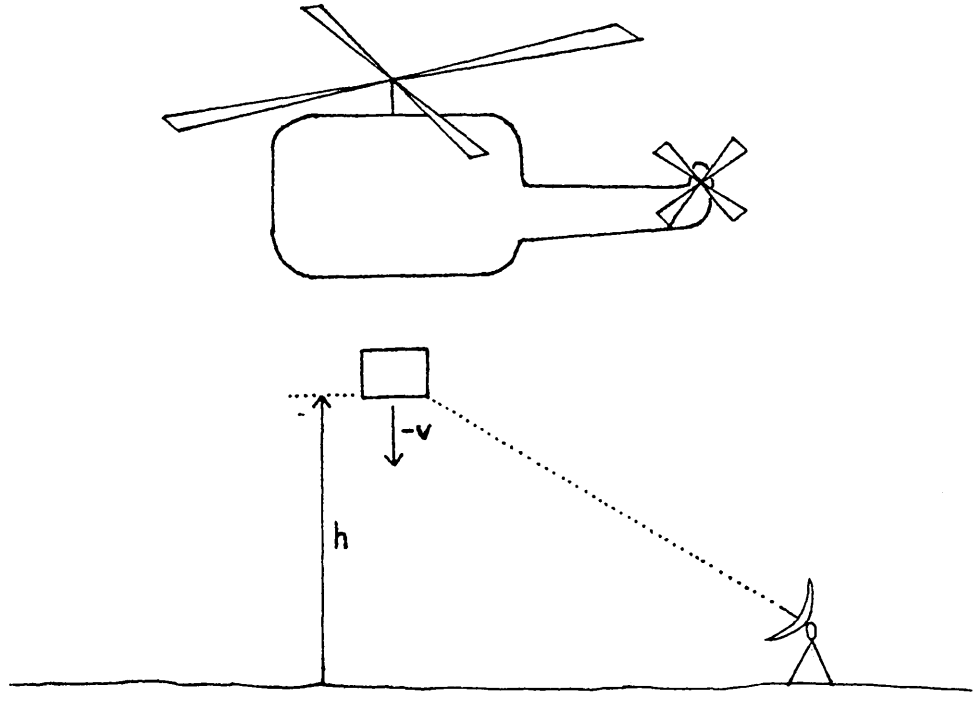


Figure A.1. Tracking a Falling Object.

The structure of the Kalman filter is shown in schematic form in Figure A.2. The filter functions by attempting to match a mathematical model of the system it is observing to the behaviour of the

system indicated by the input measurements received so far. The operation is cyclical: each time a new measurement is presented to the filter, it computes a predicted value for that input using the current internal system model (the system state model). The difference between the predicted and actual measurement (the innovation) is used to correct the current internal model so that it matches the observed behaviour of the system more closely.

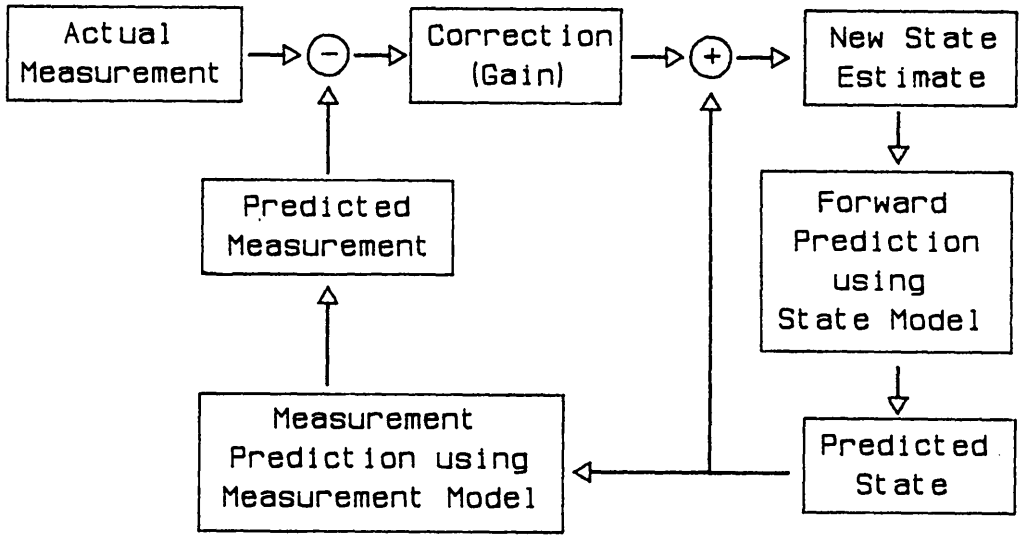


Figure A.2. Schematic Structure of a Kalman Filter.

The Kalman filter's system state model consists of three things: a vector of parameters, the state vector, that describes the current state of the system being observed; a mechanism for calculating future values of the state parameters from current values; and a mechanism for predicting the value of a measurement that would be generated by the system, using the values of its state parameters.

In terms of our example, the state parameters of the falling body are its height h and (upward) vertical velocity v, so the state vector for the filter (say $\underline{x}$) is $\begin{bmatrix} h \\ v \end{bmatrix}$. The body is falling under gravity, so $\underline{x}$ varies with time according to the equation

$$\underline{x}(t+\tau) = \phi(\tau)\underline{x}(t) + \underline{b}(\tau), \qquad (A.1)$$

where $\phi(\tau) = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix}$, $\quad \underline{b}(\tau) = g\tau \begin{bmatrix} \frac{\tau}{2} \\ 1 \end{bmatrix}$,

and g is the acceleration due to gravity. The matrix $\phi(\tau)$ which describes the evolution of the state apart from external influences is called the state transition matrix, and the vector $\underline{b}(\tau)$ is a forcing vector that models the effect of gravity. Equation (A.1) defines the state transition model for the Kalman filter application, specifying how to compute the state of the falling body at any instant given its state at some initial time.

The measurement model, which relates the measurements obtained from a system to its state parameters, is also simpe for the falling body. A measurment is generated by taking the body's height and adding a random measurement noise. If the noise is denoted ρ, then in formal terms the process is

$$\text{Predicted Measurement} = [1\ 0]\underline{x} + \rho. \qquad (A.2)$$

The matrix $[1\ 0]$ that selects the height component of the state vector is called the measurement matrix, and is generally denoted by H. (This is not connected with the h in $\underline{x}$.)

With these model definitions, the operation of the Kalman filter is as follows. Suppose that the current estimate of the system state parameter vector is $\hat{\underline{x}}(t)$, and the next measurement is expected at time $t+\tau$. The filter computes a predicted state vector $\tilde{\underline{x}}(t+\tau)$ for that time using equation (A.3):

$$\tilde{\underline{x}}(t+\tau) = \phi(\tau)\hat{\underline{x}}(t) + \underline{b}(\tau). \qquad (A.3)$$

Following the tilde-caret convention introduced in Chapter four, the caret signifies that the estimate $\hat{\underline{x}}(t)$ is based on information available at or before time t, while the tilde indicates that the prediction $\tilde{\underline{x}}(t+\tau)$ is based on information available strictly before time $t+\tau$.

Any error in the estimate $\hat{\underline{x}}(t)$ is carried forward into the predicted state vector. If the covariance of the error in $\hat{\underline{x}}(t)$ is denoted by $\hat{P}(t)$, then the error covariance of the new predicted state is given by

$$\tilde{P}(t+\tau) = \phi(\tau)\hat{P}(t)\phi(\tau)^T. \qquad (A.4)$$

The predicted value of the measurement for state vector $\tilde{x}(t+\tau)$ is $H\tilde{x}(t+\tau) + \bar{\rho}$, where $\bar{\rho}$ is the expected value (mean) of the measurement error, ρ. If the new measurement is denoted by $m(t+\tau)$, the innovation is given by $m(t+\tau) - H\tilde{x}(t+\tau) - \bar{\rho}$, and the filter adjusts its state vector, constructing a new state estimate $\underline{\hat{x}}(t+\tau)$, using

$$\underline{\hat{x}}(t+\tau) = \underline{\tilde{x}}(t+\tau) + \gamma(t+\tau)\{m(t+\tau) - H\tilde{x}(t+\tau) - \bar{\rho}\}. \qquad (A.5)$$

The matrix $\gamma(t+\tau)$ is the filter gain, and it is calculated from the variances of the predicted state $\underline{\tilde{x}}(t+\tau)$ and the measurement $m(t+\tau)$ so as to minimise the variance of the new estimate $\underline{\hat{x}}(t+\tau)$. If the measurement error variance is $R(t+\tau)$ and the prediction error variance is $\tilde{P}(t+\tau)$, then $\gamma(t+\tau)$ can be shown to have the value

$$\gamma(t+\tau) = \tilde{P}(t+\tau)H^T\{H\tilde{P}(t+\tau)H^T + R(t+\tau)\}^{-1} \qquad (A.6)$$

and the error variance of the estimate $\underline{\hat{x}}(t+\tau)$, denoted by $\underline{\hat{P}}(t+\tau)$, is then given by

$$\hat{P}(t+\tau) = \{1-\gamma(t+\tau)H\}\tilde{P}(t+\tau). \qquad (A.7)$$

Equations (A.3), (A.4), (A.5), (A.6) and (A.7), supplemented by the specification of the state transition parameters $\phi(\tau)$ and $\underline{b}(\tau)$, and the measurement model parameters H, $\bar{\rho}$ and R, constitute the definition of the Kalman filter for tracking the falling body. The only major difference between this filter and the filters used in section 4.1, for example, is that the state parameters evolve deterministically with time -- i.e. the falling body's trajectory is perfectly predictable, given accurate knowledge of its initial state.

In practice, the state of the system being tracked often cannot be perfectly predicted because it contains random variation. In the case of the falling body, for example, the trajectory might be perturbed by wind. To allow for these random effects, two more elements are added to the state transition model -- a random noise vector $q(\tau)$, called the state transition noise (or system noise or plant noise), and its covariance matrix $Q(\tau)$. The state prediction equation now becomes

$$\tilde{x}(t+\tau) = \phi(\tau)\hat{x}(t) + \underline{b}(\tau) + \overline{\underline{q}}(\tau), \qquad (A.7)$$

where $\overline{\underline{q}}(\tau)$ is the expected value (mean) of $\underline{q}(\tau)$, and the prediction error covariance computation becomes

$$\tilde{P}(t+\tau) = \phi(\tau)\hat{P}(t)\phi(\tau)^T + Q(\tau) \qquad (A.8)$$

to allow for the additional prediction error caused by the noise $\underline{q}(\tau)$.

In summary, then, the Kalman filter is defined by the five equations (suppressing the explicit time dependence):

$$\tilde{x} = \phi\hat{x} + \underline{b} + \overline{\underline{q}},$$

$$\tilde{P} = \phi\hat{P}\phi^T + Q,$$

$$\Upsilon = \tilde{P}H^T\{H\tilde{P}H^T + R\}^{-1},$$

$$\hat{x} = \tilde{x} + \Upsilon\{m - H\tilde{x} - \overline{\rho}\},$$

$$\hat{P} = \{1 - \Upsilon H\}\tilde{P},$$

and the associated definitions of the various model quantities ϕ, $\overline{\underline{q}}$, Q, $\underline{b}$ (transition model), and H, R, $\overline{\rho}$ (measurement model). The first two of the equations are used for forward prediction of the filter state; the last three are used to incorporate new measurement information into the filter state estimate.

Appendix B. The Information Averaging Filter.

The problem addressed by the Information Averaging Filter is the linear combination of a number of noisy estimates of a quantity in a manner that yields the minimal square error composite. Informally, each input estimate will contribute to the filter output in a way that depends on the information it contains.

In formal terms, the Information Averaging Filter solves this problem.

> Given N unbiassed estimates $\underline{\hat{x}}_i$ of the quantity $\underline{\xi}$, containing uncorrelated estimation errors with covariances $\hat{\pi}_i$ for estimate i, determine appropriate weight matrices β_i such that the linear combination $\underline{\hat{x}} = \sum_{i=1}^{i=N} \beta_i \underline{\hat{x}}_i$ is unbiassed and the composite error covariance $\hat{\pi}$ is as small as possible.

I shall derive the values of β_i for the case where there are two input estimates to be combined, and then state the result for N inputs. The general case follows from the two input case by an easy inductive proof which is left to the reader.

In the two estimate case the composite estimate is given by

$$\underline{\hat{x}} = \beta_1 \underline{\hat{x}}_1 + \beta_2 \underline{\hat{x}}_2 \qquad (B.1)$$

and the expected value of $\underline{\hat{x}}$ is therefore

$$\begin{aligned} E\left[\underline{\hat{x}} \right] &= \beta_1 E\left[\underline{\hat{x}}_1 \right] + \beta_2 E\left[\underline{\hat{x}}_2 \right] \\ &= (\beta_1 + \beta_2)\underline{\xi}, \end{aligned} \qquad (B.2)$$

since the input estimates are unbiassed. The composite estimate is also unbiassed, so $\beta_1+\beta_2$ must equal unity. Substituting this into (B.1),

$$\underline{\hat{x}} = \beta_1 \underline{\hat{x}}_1 + (1-\beta_1)\underline{\hat{x}}_2. \qquad (B.3)$$

Now the error covariance of the composite is

$$\hat{\pi} = E\left[(\hat{\underline{x}}-\underline{\xi})(\hat{\underline{x}}-\underline{\xi})^T \right] = \beta_1\hat{\pi}_1\beta_1^T + (1-\beta_1)\hat{\pi}_2(1-\beta_1)^T, \qquad (B.4)$$

where I have used the fact that the input estimation errors are uncorrelated. Equation (B.4) can be rearranged to give

$$\hat{\pi} = \hat{\pi}_2 - \beta_1\hat{\pi}_2 - \hat{\pi}_2\beta_1^T + \beta_1(\hat{\pi}_1+\hat{\pi}_2)\beta_1^T. \qquad (B.5)$$

The error covariance matrices are symmetric and, assuming that their sum is invertible, the value of β_i that minimises the matrix $\hat{\pi}$ is given by

$$\beta_1 = \hat{\pi}_2(\hat{\pi}_1+\hat{\pi}_2)^{-1} \qquad (B.6)$$

and the minimal value of $\hat{\pi}$ is

$$\hat{\pi}_{min} = \hat{\pi}_2 - \hat{\pi}_2(\hat{\pi}_1+\hat{\pi}_2)^{-1}\hat{\pi}_2^T. \qquad (B.7)$$

To see this, substitute $\alpha+\hat{\pi}_2(\hat{\pi}_1+\hat{\pi}_2)$ for β_1 in equation (B.5), where α is an arbitrary matrix of suitable dimension, and note that the term containing α and its transpose is a symmetric positive definite matrix.

The expressions for the weight matrix β_1 and the minimal output error covariance can be rearranged to give the more symmetrical forms:

$$\hat{\pi}_{min} = (\hat{\pi}_1^{-1} + \hat{\pi}_2^{-1})^{-1} \quad \text{and} \quad \beta_1 = \hat{\pi}_{min}\hat{\pi}_1^{-1}. \qquad (B.8)$$

The weight matrix β_2 may similarly be written

$$\beta_2 = \hat{\pi}_{min}\hat{\pi}_2^{-1} \qquad (B.9)$$

and, on substituting these values back into equation (B.1), the Information Averaging Filter for two inputs is given by

$$\hat{\underline{x}} = \hat{\pi}_{min} \sum_{i=1}^{i=2} \hat{\pi}_i^{-1}\hat{\underline{x}}_i \qquad (B.10)$$

with $\hat{\pi}_{min}$ as above.

The general form of the filter for N inputs is defined analogously, by

$$\hat{\underline{x}} = \hat{\pi}_{min} \sum_{i=1}^{i=N} \hat{\pi}_i^{-1} \hat{\underline{x}}_i, \quad \text{where} \quad \hat{\pi}_{min} = \left\{ \sum_{i=1}^{i=N} \hat{\pi}_i^{-1} \right\}^{-1}. \qquad (B.11)$$

Appendix C. Theoretical Results Presented in Chapter Four.

This appendix contains the proofs of the theorems presented in Chapter Four, and a proof of the upper bound for error covariances used in section 4.1.6. The notation used is that of Chapter 4.

The recursive computation of Algorithm A is defined by equations (4.2), (4.3) and (4.4), which are reproduced below. The α_i and β_i are gain and weight parameters (they actually vary with time, though the time subscript is omitted and the target subscript moved across); their values are, for the purposes of this appendix, arbitrary.

$$ {}_0\underline{\hat{x}}_n = \sum_{i=1}^{i=m} \beta_i \left({}_i\underline{\tilde{\xi}}_n - {}_i\underline{\hat{x}}_n \right) \quad \text{where} \quad \sum_{i=1}^{i=m} \beta_i = 1. \tag{4.2} $$

$$ {}_0\underline{\hat{\xi}}_n = {}_0\underline{\tilde{\xi}}_n + \alpha_0 \left({}_0\underline{\hat{x}}_n - {}_0\underline{\tilde{\xi}}_n \right). \tag{4.3} $$

$$ {}_i\underline{\hat{\xi}}_n = (1-\alpha_i)\,{}_i\underline{\tilde{\xi}}_n + \alpha_i \left({}_0\underline{\hat{\xi}}_n + {}_i\underline{\hat{x}}_n \right). \tag{4.4} $$

The first theorem to prove is Theorem 4.1.

Theorem 4.1. The Bias Propagation Theorem.

> Suppose that the observations of relative motion in the set $\hat{X}_n$, obtained at time t_n, are unbiassed but that the absolute state predictions in $\tilde{\Xi}_n$ all contain a (common) bias $\underline{b}_n$. Then the absolute state estimates in $\hat{\Xi}_n$ also contain bias $\underline{b}_n$.

Proof.

The two conditions of the theorem are first, that the ${}_i\underline{\hat{x}}_n$ are unbiassed:

$$ E\left[{}_i\underline{\hat{x}}_n - {}_i\underline{x}_n \right] = \underline{0} \quad \text{for } i=1..m_n; \tag{C.1} $$

and second, that the predicted absolute states contain bias $\underline{b}_n$, i.e.

$$ E\left[{}_i\underline{\tilde{\xi}}_n - {}_i\underline{\xi}_n \right] = \underline{b}_n \quad \text{for } i=0..m_n. \tag{C.2} $$

The true value of the measurement vector, ${}_i\underline{x}_n$, is given by

$$ {}_i\underline{x}_n = {}_i\underline{\xi}_n - {}_0\underline{\xi}_n \quad \text{for } i=1..m_n. \tag{C.3} $$

The observer composite state estimate is given by equation (4.2) above. Subtracting ${}_0\underline{\xi}_n$, the true observer absolute state, and taking expectations yields

$$ E\left[{}_0\underline{\hat{x}}_n - {}_0\underline{\xi}_n \right] = \sum_{i=1}^{i=m} \beta_i \, E\left[{}_i\underline{\tilde{\xi}}_n - {}_i\underline{\hat{x}}_n - {}_0\underline{\xi}_n \right], \tag{C.4} $$

where the linearity of the expectation operator and the fact that the β_i sum to unity have been used. Each expectation in the sum is equal to $\underline{b}_n$, for

$$ E\left[{}_i\underline{\tilde{\xi}}_n - {}_i\underline{\hat{x}}_n - {}_0\underline{\xi}_n \right] = E\left[{}_i\underline{\tilde{\xi}}_n - {}_0\underline{\xi}_n \right] - E\left[{}_i\underline{x}_n \right] $$

(because ${}_i\underline{\hat{x}}_n$ is unbiassed)

$$ = E\left[{}_i\underline{\tilde{\xi}}_n - {}_0\underline{\xi}_n \right] - E\left[{}_i\underline{\xi}_n - {}_0\underline{\xi}_n \right] $$

(using (C.3))

$$ = E\left[{}_i\underline{\tilde{\xi}}_n - {}_i\underline{\xi}_n \right] = \underline{b}_n \quad \text{by hypothesis.} \tag{C.5} $$

Since the β_i sum to unity, it follows that

$$ E\left[{}_0\underline{\hat{x}}_n - {}_0\underline{\xi}_n \right] = \underline{b}_n. \tag{C.6} $$

Subtracting the true observer absolute state from both sides of equation (4.3) and taking expectations gives

$$ E\left[{}_0\underline{\hat{\xi}}_n - {}_0\underline{\xi}_n \right] = (1-\alpha_0) E\left[{}_0\underline{\tilde{\xi}}_n - {}_0\underline{\xi}_n \right] + \alpha_0 E\left[{}_0\underline{\hat{x}}_n - {}_0\underline{\xi}_n \right]. \tag{C.7} $$

By hypothesis the first expectation is equal to $\underline{b}_n$ and the result just proved gives the second expectation equal to $\underline{b}_n$. Therefore

$$ E\left[{}_0\underline{\hat{\xi}}_n - {}_0\underline{\xi}_n \right] = \underline{b}_n. \tag{C.8} $$

Subtracting the true target state from equation (4.4) and taking expectations,

$$ E\left[{}_i\underline{\hat{\xi}}_n - {}_i\underline{\xi}_n \right] = (1-\alpha_i) E\left[{}_i\underline{\tilde{\xi}}_n - {}_i\underline{\xi}_n \right] + \alpha_i E\left[{}_0\underline{\hat{\xi}}_n + {}_i\underline{\hat{x}}_n - {}_i\underline{\xi}_n \right]. \tag{C.9} $$

Now the first expectation is equal to $\underline{b}_n$ by hypothesis, and

$$E\left[{}_0\hat{\underline{\xi}}_n + {}_i\hat{\underline{x}}_n - {}_i\underline{\xi}_n \right] = E\left[{}_0\hat{\underline{\xi}}_n - {}_0\underline{\xi}_n \right] + E\left[{}_i\hat{\underline{x}}_n - {}_i\underline{\xi}_n + {}_0\underline{\xi}_n \right]. \qquad (C.10)$$

In equation (C.10) the first expectation on the right has just been shown to equal $\underline{b}_n$ and the second is zero by equation (C.3) and the hypothesis that ${}_i\hat{\underline{x}}_n$ is unbiassed. Thus each vector ${}_i\hat{\underline{\xi}}_n$ for $i \varepsilon 1..m_n$ contains a bias $\underline{b}_n$ as required.

QED.

The proof of Theorem 4.2 is simpler.

Theorem 4.2. The Bias Prediction Theorem.

Suppose that an estimated absolute state at time t_m contains a bias $\underline{b}_m$. Then the corresponding predicted absolute state for time t_n contains a bias $\underline{b}_n$ where

$$\underline{b}_n = \phi(t_m, t_n)\underline{b}_m. \qquad (4.13)$$

Proof.

Using the forward prediction equations for the true state (4.5) and the state estimate (4.6),

$$\begin{aligned} E\left[{}_i\tilde{\underline{\xi}}_n - {}_i\underline{\xi}_n \right] &= E\left[\phi(t_m, t_n)\,{}_i\hat{\underline{\xi}}_m + E\left[\underline{q}_{mn} \right] - \phi(t_m, t_n)\,{}_i\underline{\xi}_m - \underline{q}_{mn} \right] \\ &= \phi(t_m, t_n) E\left[{}_i\hat{\underline{\xi}}_m - {}_i\underline{\xi}_m \right] \quad \text{by linearity} \\ &= \phi(t_m, t_n)\underline{b}_m \quad \text{using the hypothesis.} \qquad (C.11) \end{aligned}$$

QED.

The proof of Theorem 4.3 proceeds by induction over n in the sequence of times $\{t_n\}$ at which sets of input measurements arrive.

Theorem 4.3. The Initialisation Theorem.

Suppose that an estimate of the observer absolute state is available for some time t_0 and that it contains bias $\underline{b}_0$. Suppose further that no echo sources are known at t_0 and that the state transition matrix depends only on the time (and not on the source). Then, using the initialisation procedure (4.12), the estimated or predicted absolute state vectors for all echo

sources (and the observer) known at time $t_n \geq t_0$ contain a common bias $\underline{b}_n$ given by

$$\underline{b}_n = \phi(t_0, t_n)\underline{b}_0. \tag{4.8}$$

Proof.

First, we note that the theorem is trivially true at time t_0.

Now, suppose for induction that the theorem holds for all t_m in the sequence of times provided that $t_m < t_n$. The output set of the algorithm, $\hat{\Xi}_n$, contains two classes of vector: those being updated at time t_n, for which a predicted state is present in $\tilde{\Xi}_n$; and those that are newly initialised at time t_n. For the former class it is enough to prove that all the state predictions in $\tilde{\Xi}_n$ share a common bias $\underline{b}_n$ of the form required by this theorem, since Theorem 4.1 will then apply. For the latter class, consider

$$\begin{aligned} E\left[{}_i\hat{\underline{\xi}}_n - {}_i\underline{\xi}_n \right] &= E\left[{}_0\hat{\underline{\xi}}_n + {}_i\hat{\underline{x}}_n - {}_i\underline{\xi}_n \right] \\ &= E\left[{}_0\hat{\underline{\xi}}_n - {}_0\underline{\xi}_n \right] + E\left[{}_i\hat{\underline{x}}_n - {}_i\underline{\xi}_n + {}_0\underline{\xi}_n \right]. \end{aligned} \tag{C.12}$$

The latter expectation in (C.12) is zero by hypothesis. The former is $\underline{b}_n$, since the observer state vector is of the previous class of state vectors in $\hat{\Xi}_n$.

There remains the proof that the predicted state vectors in $\tilde{\Xi}_n$ all contain the correct bias $\underline{b}_n$. For this proof we need a simple lemma.

Lemma 1. The Exponential Property of $\phi(t_m, t_n)$.

For any three times, t_m, t_k and t_n,

$$\phi(t_m, t_n) = \phi(t_k, t_n)\phi(t_m, t_k).$$

Proof.

Let $\underline{\xi}_m$, $\underline{\xi}_k$ and $\underline{\xi}_n$ be appropriate state vectors for times t_m, t_k and t_n respectively. Then, using the state forward prediction equation (4.5) three times,

$$\underline{\xi}_k = \phi(t_m,t_k)\underline{\xi}_m + \underline{q}_{mk}, \tag{C.13}$$

$$\underline{\xi}_n = \phi(t_k,t_n)\underline{\xi}_k + \underline{q}_{kn}, \text{ and} \tag{C.14}$$

$$\underline{\xi}_n = \phi(t_m,t_n)\underline{\xi}_m + \underline{q}_{mn}. \tag{C.15}$$

Comparing the two ways of computing $\underline{\xi}_n$ from $\underline{\xi}_m$ (directly via equation (C.15), or by using (C.13) and (C.14) in sequence) establishes the result.

QED.

Proof of Theorem 4.3, continued.

With this lemma established, the proof proceeds as follows. Each vector $_i\underline{\tilde{\xi}}_n$ in the set $\tilde{\Xi}_n$ is a forward prediction of a state vector $_i\underline{\tilde{\xi}}_m$ estimated at some previous time t_m. For that vector the theorem holds, by the inductive hypothesis, so it contains a bias

$$\underline{b}_m = \phi(t_0,t_m)\underline{b}_0. \tag{C.16}$$

Applying Theorem 4.2 to this forward prediction step, the predicted state vector has a bias

$$\underline{b}_n = \phi(t_m,t_n)\underline{b}_m = \phi(t_m,t_n)\phi(t_0,t_m)\underline{b}_0, \tag{C.17}$$

which by lemma 1 is just $\phi(t_0,t_n)\underline{b}_n$. Now that bias vector is independent of the target number i, and therefore appears in common in all the states in the prediction set $\tilde{\Xi}_n$, which was the required intermediate result. Theorem 4.3 is therefore established.

QED.

The final result to be proved in this appendix is that if $\underline{a}$ and $\underline{b}$ are random variables with zero mean and error covariances A and B respectively, then the error covariance of their sum (or difference)

is bounded by $2(A+B)$.

Proof.

First, define the natural partial ordering on dimensionally compatible square matrices:

$$A \leq B \text{ if and only if } \underline{x}^T(B-A)\underline{x} \geq 0 \text{ for all } \underline{x}. \tag{C.18}$$

Thus $A \leq B$ if $B-A$ is a non-negative definite matrix. Under this ordering we aim to show that, for $\lambda = \pm 1$,

$$\mathbf{cov}\left[\underline{a}+\lambda\underline{b} \right] \leq 2\{A+\lambda^2 B\}. \tag{C.19}$$

Note that $A = \mathbf{cov}\left[\underline{a} \right]$ and $B = \mathbf{cov}\left[\underline{b} \right]$.

Now the covariance matrix for a vector is defined by

$$\mathbf{cov}\left[\underline{a} \right] = \mathbf{E}\left[\underline{a}\underline{a}^T \right] - \mathbf{E}\left[\underline{a} \right]\mathbf{E}\left[\underline{a}^T \right] = \mathbf{E}\left[\delta\underline{a}\ \delta\underline{a}^T \right], \tag{C.20}$$

where $\delta\underline{a}$ is $\underline{a}-\mathbf{E}\left[\underline{a} \right]$. Thus,

$$\mathbf{cov}\left[\underline{a}+\lambda\underline{b} \right] = A + \lambda^2 B + \mathbf{E}\left[\delta\underline{a}(\lambda\delta\underline{b})^T \right] + \mathbf{E}\left[(\lambda\delta\underline{b})\delta\underline{a}^T \right]. \tag{C.21}$$

Consider, for an arbitrary vector $\underline{x}$, the product

$$\underline{x}^T\mathbf{cov}\left[\underline{a}+\lambda\underline{b} \right]\underline{x} = \underline{x}^T(A + \lambda^2 B)\underline{x} + 2\mathbf{E}\left[\{\delta\underline{a}.\underline{x}\}\{\lambda\delta\underline{b}.\underline{x}\} \right], \tag{C.22}$$

where the associativity of matrix multiplication and the linearity of expectations have been used. Now for any real numbers α and β, $2|\alpha\beta| \leq \alpha^2 + \beta^2$, with equality when $\alpha = \pm\beta$. In this inequality choose $\alpha = \delta\underline{a}.\underline{x}$ and $\beta = \lambda\delta\underline{b}.\underline{x}$. Then

$$\begin{aligned} 2\mathbf{E}\left[\alpha\beta \right] &\leq 2\mathbf{E}\left[|\alpha\beta| \right] \\ &\leq \mathbf{E}\left[\alpha^2 + \beta^2 \right] \\ &\leq \underline{x}^T A\underline{x} + \lambda^2\underline{x}^T B\underline{x}. \end{aligned} \tag{C.23}$$

and so

$$\mathbf{cov}\left[\underline{a} + \lambda\underline{b} \right] \leq 2(A + \lambda^2 B) \tag{C.24}$$

as required (the result is actually true for all real λ, but we require it only for $\lambda = \pm 1$).

QED.

Appendix D. Statistical Results for Chapter Five.

This appendix establishes the results given in section 5.3.3 concerning the bias properties of the angular velocity estimator used there (equation (5.4)):

$$\hat{\underline{\omega}} = \frac{\underline{r}\times\underline{v}}{\underline{r}.\underline{r}}. \qquad (D.1)$$

Here $\underline{v}$ is a target velocity and $\underline{r}$ is the corresponding target position. These vectors are considered to be three-dimensional for notational convenience, the extra component being a zero. The angular velocity vectors are also three-dimensional but have their first two components always zero.

The first result to be proved is this. If the vectors $\underline{v}$ and $\underline{r}$ are independent Gaussian random vectors with means $\underline{v}_e$ and $\underline{r}_e$ respectively, and the position noise (in $\underline{r}$) is isotropic with variance σ^2, then

$$E\left[\underline{\omega}_e - \hat{\underline{\omega}}\right] = \underline{\omega}_e \exp(-M), \text{ where } \underline{\omega}_e = \frac{\underline{r}_e\times\underline{v}_e}{\underline{r}_e.\underline{r}_e} \text{ and } M = \frac{\underline{r}_e.\underline{r}_e}{2\sigma^2}. \qquad (D.2)$$

Proof.

Since the noises in the vectors $\underline{r}$ and $\underline{v}$ are independent and expectation is linear, it follows immediately that

$$E\left[\hat{\underline{\omega}}\right] = E\left[\frac{\underline{r}}{\underline{r}.\underline{r}}\right] \times E\left[\underline{v}\right] = E\left[\frac{\underline{r}}{\underline{r}.\underline{r}}\right] \times \underline{v}_e. \qquad (D.3)$$

The outstanding expectation is a double integral over the possible values of $\underline{r}$:

$$E\left[\frac{\underline{r}}{\underline{r}.\underline{r}}\right] = \iint \frac{\underline{r}}{\underline{r}.\underline{r}}\, pdf(\underline{r})dS. \qquad (D.4)$$

Given the fact that the position noise is isotropic, the probability density function of $\underline{r}$ is given by

$$\text{pdf}(\underline{r}) = \frac{1}{2\pi\sigma^2}\,\exp\left(-\frac{\delta\underline{r}.\delta\underline{r}}{2\sigma^2}\right), \quad \text{where } \delta\underline{r} = \underline{r}-\underline{r}_e. \tag{D.5}$$

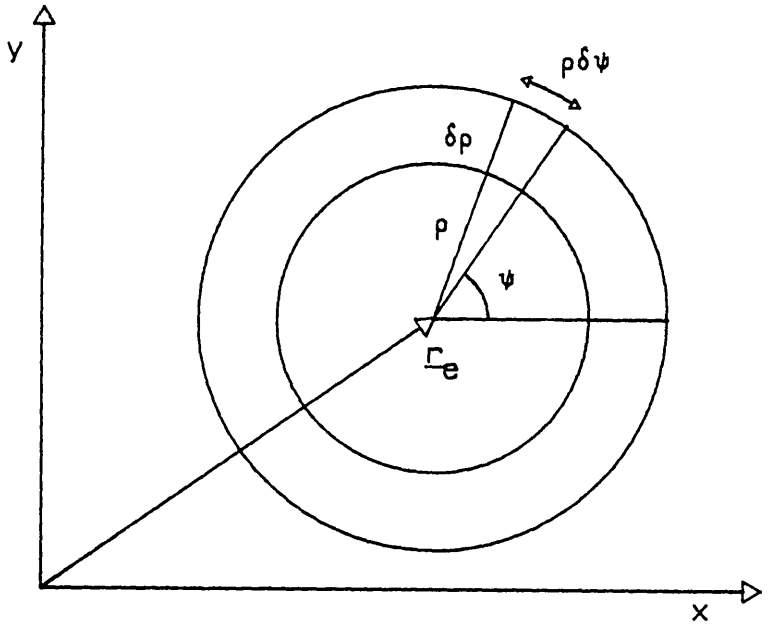


Figure D.1. Transforming the Integral to Polar Coordinates.

Substituting (D.5) into the double integral (D.4) and performing the area integration using polar coordinates (ρ,ψ) centred on the expected position $\underline{r}_e$ gives

$$E\left[\frac{\underline{r}}{\underline{r}.\underline{r}}\right] = \frac{1}{\sigma^2}\int_0^\infty I(\rho)\,\rho\exp\left(-\frac{\rho^2}{2\sigma^2}\right)d\rho, \quad \text{where} \tag{D.6}$$

$$I(\rho) = \frac{1}{2\pi}\int_0^{2\pi}\frac{\underline{r}}{\underline{r}.\underline{r}}\,d\psi. \tag{D.7}$$

To evaluate the integral (D.7), consider the plane of values of $\underline{r}$ to be a plane of complex numbers. There is a natural isomorphism between the 2-vectors and the complex numbers, so let the image of $\underline{r}_e$ under that isomorphism be the complex number m and then the image of any value of $\underline{r}$ in the integral (D.7) is $m+\rho z$ for some complex number z on the unit circle. Making this substitution and transforming the integral into a contour integral round the unit circle,

$$I(\rho) \approx \frac{1}{2\pi i} \oint \frac{dz}{\rho + \bar{m}z}. \qquad (D.8)$$

(The bar on $\bar{m}$ denotes complex conjugation, and the symbol $\approx$ denotes equivalence under the natural isomorphism.) Using the theory of residues, this integral is equal to the net residue of its integrand within the contour. The value of the integral is

$$I(\rho) \approx \begin{cases} \frac{1}{m} & \text{for } \rho < |m| \\ 0 & \text{for } \rho > |m| \end{cases} \qquad (D.9)$$

and, on substituting these values into the integral (D.6), noting that the image of $1/\bar{m}$ under the natural isomorphism is $\frac{\underline{r}_e}{\underline{r}_e \cdot \underline{r}_e}$, the expectation of (D.6) becomes

$$E\left[\frac{\underline{r}}{\underline{r} \cdot \underline{r}} \right] = \frac{\underline{r}_e}{\underline{r}_e \cdot \underline{r}_e} \frac{1}{\sigma^2} \int_0^{|\underline{r}_e|} \rho \exp\left(-\frac{\rho^2}{2\sigma^2}\right) d\rho. \qquad (D.10)$$

The required expression (D.2) follows at once on evaluating this integral.

QED.

The second result relaxes the assumptions made above, allowing correlation between the velocity and position noises though retaining the Gaussian distributions and the isotropic position noise. Given these assumptions the extra bias contribution to $E\left[\hat{\underline{\omega}} - \underline{\omega}_e \right]$ from the noise correlation is bounded in magnitude by a constant multiple of $\frac{1}{M}$, where M is as defined above in (D.2).

Proof.

In this more complicated case, the expected value of $\hat{\underline{\omega}}$ is given by the four dimensional volume integral over all values of $\underline{r}$ and $\underline{v}$:

$$E\left[\hat{\underline{\omega}}\right] = \iiint \frac{\underline{r}}{\underline{r}\cdot\underline{r}} \times \underline{v}\ \mathrm{pdf}(\underline{r},\underline{v})dV, \tag{D.12}$$

and the joint probability density function is given by

$$\mathrm{pdf}(\underline{r},\underline{v}) = \frac{\mathbf{det}P}{4\pi^2}\ \mathbf{exp}\left(-\frac{1}{2}Q(\delta\underline{r},\delta\underline{v})\right), \quad \text{where} \tag{D.13}$$

$$Q(\delta\underline{r},\delta\underline{v}) = \delta\underline{r}^T P_{rr}\delta\underline{r} + \delta\underline{r}^T P_{rv}\delta\underline{v} + \delta\underline{v}^T P_{rv}^T\delta\underline{r} + \delta\underline{v}^T P_{vv}\delta\underline{v}.$$

Here, the vector $\delta\underline{v}$ is $\underline{v}-\underline{v}_e$, and the matrices P_{rr}, P_{rv} and P_{vv} are blocks of the joint error covariance matrix P:

$$P = \begin{bmatrix} P_{rr} & P_{rv} \\ P_{rv}^T & P_{vv} \end{bmatrix}. \tag{D.14}$$

By making the substitution $\underline{v} = \underline{y} + A\underline{r}$ for a suitably chosen matrix A, the quadratic term in (D.13) can be separated into two terms, and the joint probability density function split into two parts:

$$\mathrm{pdf}(\underline{r},\underline{v}) = \left\{\frac{\mathbf{det}R}{2\pi}\ \mathbf{exp}\left(-\frac{\delta\underline{r}^T R\delta\underline{r}}{2}\right)\right\}\left\{\frac{\mathbf{det}Y}{2\pi}\ \mathbf{exp}\left(-\frac{\delta\underline{y}^T Y\delta\underline{y}}{2}\right)\right\}. \tag{D.15}$$

In (D.15), R and Y are the inverse covariance matrices for $\delta\underline{r}$ and $\delta\underline{y}$ respectively.

Substituting for $\underline{v}$ in the volume integral (D.12) and using (D.14) gives

$$E\left[\hat{\underline{\omega}}\right] = \iiint \left\{\frac{\underline{r}}{\underline{r}\cdot\underline{r}} \times \underline{y} + \frac{\underline{r}}{\underline{r}\cdot\underline{r}} \times A\underline{r}\right\} \mathrm{pdf}(\underline{r})\ \mathrm{pdf}(\underline{y})\ dS_r dS_y. \tag{D.16}$$

Now the double integral over $\underline{y}$ may be done at once, since the probability density function of $\underline{y}$ integrates to one and $\underline{y}\ \mathrm{pdf}(\underline{y})$ integrates to $\underline{y}_e$, the expected value of $\underline{y}$. The two probability densities are just the two parts of equation (D.14). Adding and subtracting $\frac{\underline{r}}{\underline{r}\cdot\underline{r}} \times A\underline{r}_e$ and performing the integration over $\underline{y}$, the volume integral becomes

$$E\left[\hat{\underline{\omega}}\right] = E\left[\frac{\underline{r}}{\underline{r}\cdot\underline{r}}\right] \times \underline{v}_e + \iint \frac{\underline{r}}{\underline{r}\cdot\underline{r}} \times A\delta\underline{r}\ \mathrm{pdf}(\underline{r})dS_r. \tag{D.17}$$

The expectation on the left of the right hand side in (D.17) is known

from the first result to be of order e^{-M}; the double integral is the contribution to the estimate from the coupled noise. It is therefore sufficient to consider this latter term, which I shall denote by I_{cor}.

Since the position noise is isotropic, R is a diagonal matrix. Substituting for it explicitly and converting the resulting double integral into polar coordinate form, analogously to (D.6) and (D.7), gives

$$I_{cor} = \frac{1}{\sigma^2} \int_0^\infty I(\rho)\, \rho \exp\left(-\frac{\rho^2}{2\sigma^2}\right) d\rho, \quad \text{where} \tag{D.18}$$

$$I(\rho) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\underline{r}}{\underline{r}.\underline{r}} \times A\delta\underline{r}\, d\psi. \tag{D.19}$$

As before, $\underline{r}$ in this integral takes values on the circumference of the circle centred at $\underline{r}_e$ with radius ρ.

To transform the integral in (D.19) into a contour integral, note that the image, under the natural isomorphism from 2-vectors to complex numbers, of the matrix product $A\delta\underline{r}$ is given by the linear function $\rho(\alpha z+\beta\bar{z})$ for suitably chosen complex constants α and β. The cross-product term in (D.19), which is a vector with a single non-zero component, maps into the imaginary part of the complex number $\rho(\bar{m}+\rho\bar{z})(\alpha z+\beta\bar{z})$, where m is the image of $\underline{r}_e$ and z takes values on the unit circle. On making these substitutions and eliminating the non-analytic operations (conjugation of z and the 'imaginary part' operation) the integral becomes

$$I(\rho) \approx \frac{-\rho}{4\pi} \oint \frac{\alpha m^2+\beta\rho^2}{m^2(m+\rho z)} + \frac{\rho\bar{\beta}}{\bar{m}(\rho+\bar{m}z)} - \frac{\alpha\bar{m}+\beta\rho^2}{\rho\bar{m}z} + \ldots\, dz, \tag{D.20}$$

where the terms represented "..." contribute no residue in or on the contour of integration. The integrand has a pole at $z = 0$ and either a pole at $z = -\rho/\bar{m}$, when $\rho<|m|$, or a pole at $z = -m/\rho$, when $\rho>|m|$. The residue from the latter pole cancels out that from the pole at zero, so the integral vanishes for $\rho>|m|$; for $\rho<|m|$, however, it has

the value

$$I(\rho) \approx \rho^2 \, \mathbf{Im}\{ \frac{\bar{\beta}}{\bar{m}^2} \}, \qquad (D.21)$$

where the function **Im**{ } returns the imaginary part of its argument (removing the i). Using the natural isomorphism to convert back to vector form, and recalling that the imaginary part is the image of a vector cross-product and that the reciprocal of $\bar{m}$ maps into $\frac{\underline{r}_e}{\underline{r}_e \cdot \underline{r}_e}$,

$$\mathbf{Im}\{ \frac{\bar{\beta}}{\bar{m}^2} \} \approx \frac{\underline{r}_e}{\underline{r}_e \cdot \underline{r}_e} \times \frac{\underline{B}\underline{r}_e}{\underline{r}_e \cdot \underline{r}_e} \qquad (D.22)$$

for a matrix B related to the original transformation matrix A. Then

$$I_{cor} = \frac{\underline{r}_e}{\underline{r}_e \cdot \underline{r}_e} \times \frac{\underline{B}\underline{r}_e}{\underline{r}_e \cdot \underline{r}_e} \; \frac{1}{\sigma^2} \int_0^{|\underline{r}_e|} \rho^2 \, \rho \exp(-\frac{\rho^2}{2\sigma^2}) \, d\rho$$

$$= \frac{\underline{r}_e \times \underline{B}\underline{r}_e}{\underline{r}_e \cdot \underline{r}_e} \; \frac{1}{M}\{1 - (M+1)e^{-M}\}. \qquad (D.23)$$

It may be shown, on substituting for the matrix B, that the magnitude of the cross-product term $\frac{\underline{r}_e \times \underline{B}\underline{r}_e}{\underline{r}_e \cdot \underline{r}_e}$ is independent of the length of $\underline{r}_e$, and is in fact half the difference of the leading diagonal components of the original matrix A. Thus the bias integral I_{cor} is of order $\frac{1}{M}$ and the result is established.

QED.

Bibliography.

Alspach, D L; 1975.

"A Gaussian Sum Approach to the Multitarget Identification-Tracking Problem". Automatica **11**, 285-296.

Andreas, R D, Hostetler, L D, and Beckman, R C; 1978.

"Continuous Kalman Updating of an Inertial Navigation System using Terrain Measurements". Proc. Nat. Aerosp. Electron. Conf. Dayton, OH, 1263-1270.

Bard, Y; 1974.

"Non-linear Parameter Estimation". Academic Press.

Barrow, H G, and Tenenbaum, J M; 1980.

"Reconstructing Smooth Surfaces from Partial, Noisy Information". SRI International AI Centre Tech. Note 222.

Bar-Shalom, Y; 1974.

"Extension of the Probabilistic Data Association Filter to Multitarget Environments". Proc. 5th Symposium on Non-linear Estimation Theory and its Applications, San Diego, CA, 16-21.

Bar-Shalom, Y; 1978.

"Tracking Methods in a Multitarget Environment". IEEE Trans. Automatic Control **AC-23**, 618-626.

Bar-Shalom, Y, and Birmiwal, K; 1982.

"Variable Dimension Filter for Manoeuvring Target Tracking". IEEE Trans. Aerospace and Electronic Systems **AES-18**, 621-628.

Bar-Shalom, Y, and Jaffer, A G; 1972.

"Adaptive Non-Linear Filtering for Tracking with Measurements of Uncertain Origin". Proc. IEEE Conf. Decision and Control, New Orleans, 243-247.

Bar-Shalom, Y, and Tse, E; 1975.

"Tracking in a Cluttered Environment with Probabilistic Data Association". Automatica **11**, 451-460.

Bauzil, G, Briot, M, and Riges, P; 1981.

"A Navigation Subsystem using Ultrasonic Sensors for the Mobile Robot HILAIRE". Proc. First Int. Conf. Robot Vision and Sensory Controls, Stratford-upon-Avon, 47-58.

Bozic, S M; 1979.

"Digital and Kalman Filtering". Edward Arnold Ltd., Part 2.

Brady, M, and Grimson, W E L; 1981.

"The Perception of Subjective Surfaces. MIT AI Memo no. 666.

Brooks, R A; 1981.

"Symbolic Reasoning among Three-dimensional Models and Two-dimensional Images". Artificial Intelligence **17**, 285-348.

Buchner, M R; 1977.

"A Multistatic Track Filter with Optimal Measurement Selection". Proc. RADAR-77 IEE Int. Conf., London, 72-75.

Clocksin, W F; 1978.

"Determining the Orientation of Surfaces from Optical Flow". Proc. Conf. AISB/GI, Hamburg, 93-102.

Demetry, J S, and Titus, H A; 1968.

"Adaptive Tracking of Manoeuvring Targets". IEEE Trans. Automatic Control **AC-13**, 749-750.

Duck, G M, Goodson, A D, and Griffins, J W R; 1984.

"Seavision -- a New Sector Scanning Sonar". IEE Colloquium on Underwater Navigation, Professional Group E15.

Dunbar, R M, and Holmes, R T; 1978.

"Advances in Remotely Controlled Submersible Research". Oceanology International.

Fisher, R B; 1983.

"Using Surfaces and Object Models to Recognise Partially Obscured Objects". Proc. 8th IJCAI, Karlsruhe, 989-995.

Fitzgerald, R J; 1981.

"Simple Tracking Filters: Closed-Form Solutions". IEEE Trans. Aerospace and Electronic Systems **AES-17**, 781-785.

Fortmann, T E, Bar-Shalom, Y, and Scheffé, M; 1980.

"Multi-Target Tracking using Joint Probabilistic Data Association". Proc. IEEE Conf. Descision and Control, Albuquerque, NM, 807-812.

Fortmann, T E, Bar-Shalom, Y, and Scheffé, M; 1983.

"Sonar Tracking of Multiple Targets using Joint Probabilistic Data Association". IEEE J. Oceanic Engineering **OE-8** no. 3, 173-183.

Freedman, A; 1962.

"The Formation of Acoustic Echoes in Fluids". PhD thesis, London University (External).

Friedland, B; 1969.

"Treatment of Bias in Recursive Filtering". IEEE Trans. Automatic Control **AC-14**, 359-367.

Gennery, D B; 1977.

"A Stereo Vision System for Autonomous Vehicles". Proc. 5th IJCAI, Cambridge, MA, 576-582.

Gupta, S N, and Ahn, S M; 1983.

"Closed-Form Solutions of Target-Tracking Filters with Discrete Measurements". IEEE Trans. Aerospace and Electronic Systems **AES-19**, 532-538.

Hannah, M J; 1980.

"Bootstrap Stereo". Proc. AAAI, Stanford CA, 38-40.

Helmholtz, H von; 1925.

"Treatise on Physiological Optics". Dover Publications Inc, New York. Volume 3, (ed. Southall, J P C), 295.

Holmes, J E; 1977.

"The Development of Algorithms for the Formation and Updating of Tracks". Proc. RADAR-77 IEE Int. Conf., London, 81-85.

Horn, B K P, and Ikeuchi, K; 1979.

"Numerical Shape from Shading and its Applications". MIT AI Lab. Working Paper 196.

Hostetler, L D, and Andreas, R D; 1983.

"Non-Linear Kalman Filter Techniques for Terrain-Aided Navigation. IEEE Trans. Automatic Control **AC-28**, 315-323.

Jaffer, A G, and Bar-Shalom, Y; 1972.

"On Optimal Tracking in Multiple Target Environments". Proc. 3rd Symposium on Nonlinear Estimation Theory and its Applications, San Diego CA, 112-117.

Kalman, R A; 1960.

"A New Approach to Linear Filtering and Prediction Problems". Trans. ASME J. Basic Engineering **82 D**, 35-46.

Kanade, T; 1979.

"Recovery of the Three-Dimensional Shape of an Object from a Single View". Carnegie-Mellon University Computer Science Report CMU-CS-79-153.

Kenefic, R J; 1981.

"Optimum Tracking of a Manoeuvring Target in Clutter". IEEE Trans. Automatic Control **AC-26**, 750-753.

Koenderink, J J, and van Doorn, A J; 1977.

"How an Ambulant Observer can Construct a Model of the Environment from the Geometrical Structure of the Visual Inflow". Kybernetik, ed. G. Hauske & E. Butenandt, 224-247.

Lane, D M; 1984.

"An Investigation of the Intelligent Guidance and Control Systems for an Autonomous Submersible System -- Interim Report". Underwater Technology Group, Heriot-Watt University, Edinburgh.

Larcombe, M H E; 1981.

Personal communication.

Longuet-Higgins, H C, and Prazdny, K F; 1980.

"The Interpretation of a Moving Retinal Image". Proc. Roy. Soc. London **B208**, 385-397.

Marks, B L; 1961.

"Adjustment Rules for Automatic Tracking". RAE Technical Note Math. 79.

Marr, D; 1982.

"Vision". W. H. Freeman & Co, San Francisco, CA.

Mattin, R B; 1982.

"Understanding Kalman Filters -- How to Extract the Maximum Information from Imperfect Measurement". Aerospace Dynamics no. **7**, 8-18.

Mealy, G L, and Tang, W; 1983.

"Application of Multiple Model Estimation to a Recursive Terrain Height Correlation System". IEEE Trans. Automatic Control **AC-28**, 323-331.

Miller, R J; 1981.

"Multiple Option Radar Tracking of a Single Target". The Marconi Review **44**, 31-56.

Moose, R L; 1975.

"An Adaptive State Estimation Solution to the Manoeuvring Target Problem". IEEE Trans. Automatic Control **AC-20**, 359-362.

Moravec, H P; 1980.

"Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover". Stanford Computer Science Report STAN-CS-80-813.

Moravec, H P; 1979.

"Visual Mapping by a Robot Rover". Proc. 6th IJCAI, Tokyo, 589-601.

Morefield, C L; 1977.

"Application of 0-1 Integer Programming to Multitarget Tracking Problems". IEEE Trans. Automatic Control **AC-22**, 302-312.

Morley, A R, and Wilsdon, A S; 1977.

"Multiradar Tracking in a Multisite Environment". Proc. RADAR-77 IEE Int. Conf., London, 66-71.

Nagel, H H; 1983.

"Constraints for the Estimation of Displacement Vector Fields from Image Sequences". Proc. 8th IJCAI, Karlsruhe, 945-951.

Orfanidis, S J; 1982.

"An Exact Solution of the Time-Invariant Discrete Kalman Filter". IEEE Trans. Automatic Control **AC-27**, 240-242.

Pardini, S, and Grasso, G; 1973.

"Data Smoothing in the Selenia Anticollision System". Riv. Tech. Selenia **1**, 23-31.

Popplestone, R J, and Ambler, P; 1977.

"Forming Body Models from Range Data". Research Report no. 46, Artificial Intelligence Department, Edinburgh University.

Prazdny, K F; 1980.

"Ego-Motion and Relative Depth Map from Optical Flow". Biological Cybernetics **36**, 87-102.

Reed, C G, and Hogan, J J; 1979.

"Range Correlation Guidance for Cruise Missiles". IEEE Trans. Aerospace and Electronic Systems **AES-15**, 547-554.

Reid, D B; 1979.

"An Algorithm for Tracking Multiple Targets". IEEE Trans. Automatic Control **AC-24**, 843-854.

Rieger, J H, and Lawton, D T; 1983.

"Sensor Motion and Relative Depth from Difference Fields of Optic Flow". Proc. 8th IJCAI, Karlsruhe, 1027-1031.

Russell, G T, and Bugge, J; 1981.

"Adaptive Estimator for Automatic Guidance of an Unmanned Submersible". IEE Proc. **128 D** no. 5, 223-226.

Schafer, G; 1976.

"A Mathematical Theory of Evidence". Princeton University Press, Princeton, NJ.

Singer, R A; 1970.

"Estimating Optimal Tracking Filter Performance for Manned Manoeuvring Targets". IEEE Trans. Aerospace and Electronic Systems **AES-6**, 473-483.

Singer, R A, and Monzingo, R A; 1971.

"Application of Optimal Tracking to Goal-Seeking Attack Vehicles". IEEE Trans. Automatic Control **AC-16**, 63-66.

Singer, R A, and Sea, R G; 1973.

"New Results in Optimising Surveillance System Tracking and Data Correlation Performance in Dense Multitarget Environments". IEEE Trans. Automatic Control **AC-18**, 571-581.

Singer, R A, Sea, R G, and Housewright, K B; 1974.

"Derivation and Evaluation of Improved Tracking Filters for use in Dense Multi-Target Environments". IEEE Trans. Information Technology **IT-20**, 423-432.

Singer, R A, and Stein, J J; 1971.

"An Optimal Tracking Filter for Processing Sensor Data of Imprecisely Determined Origin in Surveillance Systems". Proc. IEEE Conf. Decision and Control, Miami Beach, FL; 171-175.

Sittler, R W; 1964.

"An Optimal Data Association Problem in Surveillance Theory". IEEE Trans. Military Electronics **MIL-8**, 125-139.

Smith, P, and Buechler, G; 1975.

"A Branching Algorithm for Discriminating and Tracking Multiple Objects". IEEE Trans. Automatic Control **AC-20**, 101-104.

Sutton, J L; 1979.

"Underwater Acoustic Imaging". Proc. IEEE 67 no. 4, 554-566.

Symons, M; 1982.

"The Automatic Track While Scan System used within the Searchwater Airborne Maritime Surveillance Radar". Proc. RADAR-82 IEE Int. Conf., London, 254-258.

Tenney, R R, Hebbert, R S, and Sandell, N R; 1977.

"A Tracking Filter for Manoeuvring Sources". IEEE Trans. Automatic Control **AC-22**, 246-251.

Thornton, C L, and Bierman, G J; 1978.

"Filtering and Error Analysis via the UDU T Covariance Factorization". IEEE Trans. Automatic Control **AC-23**, 901-906.

Tuncliffe, R J; 1977.

"A Simple Automatic RADAR Track Extraction System". Proc. RADAR-77 IEE Int. Conf., London, 76-80.

Ullman S; 1979.

"The Interpretation of Structure from Motion". Proc. Roy. Soc. London B203, 405-426.

Wesley, L P; 1983.

"Reasoning about Control: the Investigation of an Evidential Approach". Proc. 8th IJCAI, Karlsruhe, 203-206.

Witkin, A P, and Tenenbaum, J M; 1983.

"What is Perceptual Organisation for?" Proc. 8th IJCAI, Karlsruhe, 1023-1026.

Wittenburg, J; 1977.

"Dynamics of Systems of Rigid Bodies". Pub. B G Tenbuer, Stuttgart.

Yachida, M, Ichinose, T, and Tsuji, S; 1983.

"Model-Guided Monitoring of a Building Environment by a Mobile Robot". Proc. 8th IJCAI, Karlsruhe, 1125-1127.

Yoshimura, T, and Soeda, T; 1972.

"An Approach to the Divergence Prevention of the Extended Kalman Filter". Proc. 3rd Symposium on Nonlinear Estimation Theory and its Applications, San Diego CA, 282-285.

Bibliographic Index.