Analysis and Measurement of Motion
in 2D Medical Imaging Sequences
Exploiting Uncertain Knowledge

Frank Höwing

A submission presented in partial fulfilment of the
requirements of the University of Glamorgan/Prifysgol Morgannwg
for the degree of Doctor of Philosophy

This research programme was carried out
in collaboration with the Fachhochschule Braunschweig/Wolfenbüttel
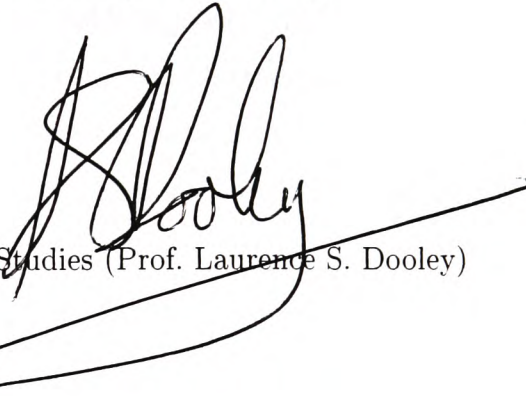
October 2001

Certificate of Research

This is to certify that, except where specific reference is made, the work described in this thesis is the result of the candidate. Neither this thesis, nor any part of it, has been presented, or is currently submitted, in candidature for any degree at any other University.

Candidate (Frank Höwing)

Director of Studies (Prof. Laurence S. Dooley)

October 2001

Date

Acknowledgments

Summary

Segmentation of moving and deformable structures in medical image processing has become an increasingly important research field in recent years. Fast and high-resolution image acquisition methods such as Magnetic Resonance (MR) imaging produce very detailed cross-sectional images of the human body. Segmentation of anatomically relevant objects is then a subsequent operation, performed in order to visualise and/or measure shapes and motions of interest. The segmentation task is usually performed by clinicians and other experts. High demand on expert time and inter- and intra-observer variability impose a clinical and scientific need of automating this process.

This thesis presents a novel approach for segmenting and tracking of anatomical objects in 2D medical image sequences, which enables quantitative studies of relevant structures even under the presence of distortions introduced by the image formation process.

The underlying premise of the work presented, is based upon the observation that a robust and precise image segmentation requires *a priori* knowledge on both image formation and the objects to be detected. Such knowledge is often vague or uncertain and may only be acquired from experts in natural-language terms.

In combining active contours with fuzzy logic a novel contour segmentation method is developed which is capable of exploiting uncertain knowledge in both syntactical and linguistic terms. Unlike other approaches the contour description is fully integrated into the segmentation process, with the additional advantage that many existing image processing operators can also smoothly be integrated into the fuzzy framework.

Specific applications addressed are motion analysis of carpal bones in MR image sequences and the analysis of the vocal tract in X-ray image sequences. Traditional solutions for both applications have been developed. The new framework is fully validated in comparison to these solutions as well as on synthetic image material.

Contents

List of Acronyms and Abbreviations

CCD	Charge Coupled Device
CCITT	Comite Consultatif International Telegraphique et Telephonique
COA	Centre Of Area
COG	Centre Of Gravity
CT	Computer Tomography
FIS	Fuzzy Inference System
FSM	Finite State Machine
MR	Magnetic Resonance
MRI	MR Imaging
PAL	Phase Alternated Line
ROI	Region Of Interest
SADT	Structured Analysis and Design Technique
$\mathcal{A}$	fuzzy set
$\mathcal{A}_c$	crisp set
$\mathbf{C}$	contour candidate vector
c	contour candidate
$\mathbf{D}$	fuzzy contour description
D_t	difference image
d_z	fuzzy segment description

E_{con}	constraint energy
E_{edge}	edge-based energy
e_{edge}	edge-based evidence
E_{ext}	external active contour energy
E_{ext}^{z}	external energy for a fuzzy rule base labelled z
E_{image}	image energy
e_{image}	image evidence
E_{int}	internal active contour energy
E_{motion}	motion-based energy
e_{motion}	motion-based evidence
e_{region}	region-based evidence
E_{snake}	active contour energy
$I(x,y)$	image intensity at the spatial co-ordinates (x,y)
l	fuzzy number
l_0	mean value of l
l_z	linguistic label for a fuzzy segment length
l_z	absolute fuzzy segment length
l_z^*	relative fuzzy segment length
M	number of candidates in C
N	number of vertices
P	initial contour
p	initial contour vertex
Q	detected contour
q	detected contour vertex
R	region of interest
s	spread of l
$\mathcal{T}$	set of linguistic terms
t	time

$\mathcal{U}$	universe of discourse for a linguistic variable
u	base variable for $\mathcal{U}$
u_i	crisp input value
u_o	crisp output value
$\mathbf{v}(s)$	contour vertex at position s (continuous representation)
w_{con}	weight for constraint energy term
w_{edge}	edge evidence weight
w_{ext}	weight for external energy term
w_{image}	weight for image energy term
w_{int}	weight for internal energy term
w_{motion}	motion evidence weight
w_{region}	region evidence weight
$\mathcal{X}$	fuzzy set
x	name of a linguistic variable
Z	number of contour segments
$\mathcal{Z}$	fuzzy set
z_z	fuzzy segment property
α	regular expression describing a contour
γ	angle
ε_{snake}	active contour energy (continuous)
ε_{int}	internal active contour energy (continuous)
ε_{image}	image energy (continuous)
ε_{con}	constraint energy (continuous)
ε_{ext}	external active contour energy (continuous)
θ	threshold
μ	membership function
σ	standard deviation
φ	angle
∇	gradient magnitude

Chapter 1

Introduction

1.1 Background

Research interest in medical image processing has grown continuously over the last three decades. Imaging devices such as digital X-ray, Computer Tomography (CT) and Magnetic Resonance Imaging (MRI) are now widespread applications of what was once considered advanced research. These devices add digital visual enhancement technologies to the everyday practice of medical experts, improving the reliability of the diagnostic process. In more recent years, computer vision research has focused on the challenge of automatically extracting anatomical structures from image sequences, leading medical applications from predominantly qualitative to quantitative analysis.

First results in image processing dating back to the early 1970s, initiated a continous cycle of improvements in computer technology, advances in image processing and algorithmic research and the medical application of more powerful computer assisted methods and tools.

With the digitisation of analogue X-ray images it became possible to enhance single images in a more advanced way than was possible with analogue signal processing techniques. The medical expert's eye was assisted by imaging devices which afforded histogram-based contrast enhancement and pseudo-colour display to name just two of

the advances. The visual diagnosis of diseases such as tumours through the detection of static anatomical structures became easier and more reliable.

Lower radiation doses and higher image resolutions were introduced by digital X-ray devices which made use of charge coupled devices (CCD) rather than traditional celluloid film. Together with improved computing capacity, visual enhancements became possible in real-time, leading to the application of new innovative devices during surgical procedures. Still more sophisticated image processing methods were required however that could automatically detect structures such as blocked blood vessels.

Another research strand, parallel to imaging and visual enhancements, provided a new and exciting dimension to medicine: the *quantitative* analysis of image sequences. Medical experts have always had a need to measure particular anatomical structures. Applications ranged from using single radiographs to measure lung areas for the diagnosis of related diseases, and the measurement of size and distance of finger bones to calculate the growth of children.

With analogue images, the measurement was performed manually by medical experts applying pen and ruler. When images became available in digital form, "electronic rulers" were used, to measure the distances between anatomically relevant landmarks, such as the ends of finger bones. Similarly areas of the lung, for example, could be measured by applying a drawing tool.

These manual techniques however, are both time-consuming and subject to individual variations of the experts' capabilities. It is therefore highly desirable to let a computer identify the relevant anatomical structures, outline their contours and measure the significant landmarks automatically. The benefits of such a system include more precise results in less time, making diagnostics more reliable and cost efficient. As measurement results will be based upon a repeatable, objective method, new areas of medical research will be possible, as well as being able also to evaluate the success of a therapy over time.

Such sophisticated techniques are particularly relevant in medical areas where the

analysis of image sequences is desired. Together with radiation-free imaging such as MRI, it will be possible to acquire mass data for research purposes and hence investigate new medical phenomena, creating statistically sound reference data for normal and pathologic cases. An example for such innovative research is the analysis of carpal bones, detailed in this thesis.

Despite the scientific and technological advances in recent years, the task of automatically identifying meaningful structures in medical images still poses a number of major challenges. Rather than classifying pixels based on local image features it is necessary to include *a priori* knowledge about the shape of anatomical structures. Furthermore, a quantitative analysis requires more precise detection results than was previously deemed sufficient for a qualitative analysis.

In addition to obtaining static images, novel imaging technologies are able to create image sequences. It is now possible to analyse two-dimensional motion (frequently referred to in the literature as 2D+t), three-dimensional structures (3D) or even three-dimensional motion (3D+t). As with two-dimensional image analysis there is a strong demand in medicine to automatically process these kinds of image sequences, example applications include the analysis of the vocal tract (2D+t), detection of brain tumours (3D) and the diagnosis of heart-related diseases (3D+t).

This demand results in an evolution in research from single images to image sequences, from static shape to motion and deformation, from pixel processing to scene interpretation and hence from signal or image processing to high-level computer vision.

There is no doubt that the long term future techniques in automatic medical image processing will reliably handle three-dimensional image sequences. These 3D and 3D+t cases, however will continue to pose very high demands on both computing resources and theoretic concepts for the foreseeable future.

High-level processing of two-dimensional image sequences on the other hand will have a big impact on a wide range of today's medical questions. Even when three-dimensional techniques become more mature and applied, 2D+t techniques will still

be attractive since they will adequately cover many medical applications and research areas. Furthermore the resources required by 2D+t techniques will be considerably less and their user interfaces inherently easier to handle than three-dimensional techniques.

1.2 Objectives

With the prediction that the analysis of two-dimensional image sequences will play a major role in future medical image processing and that related systems will coexist with three-dimensional techniques, today's systems still need to be improved to meet the requirements of quantitative medical image analysis. From an application point of view, these requirements embrace the automatic analysis and subsequent measurement of moving anatomical structures, applying a wide range of different imaging devices, such as X-ray or MRI. From a technical and scientific perspective the most crucial requirement is the precise segmentation of object outlines and their tracking over a sequence of frames, even under the presence of distortions introduced by the image formation process.

The demand to handle anatomical structures both rigid and deformable, requires a detection of a wide range of shapes which cannot be defined exactly. This degree of uncertainty in the problem definition is increased further by the fact that devices such as X-ray scanners involve an imaging process where relevant and irrelevant structures are superimposed. It was expressly to address the problem of these uncertainties, which are involved in both image formation and the characteristics of the objects to be detected, that this research project has its roots. A system was formulated with the clear objective of providing a framework for quantitative medical image analysis that would afford an exploitation of such uncertainties.

Figure 1.1 provides a conceptual insight into the system which forms the framework for the research. It involves the design of a novel feature extraction component (block 2), while providing a simple yet effective interface, decoupling the image analysis component (block 3) from a particular imaging device (block 1). An effective adaption of

this interface to a particular imaging situation is required. To achieve this, the low-level feature extraction will be able to exploit the many operators that exist as a result of traditional image processing research.
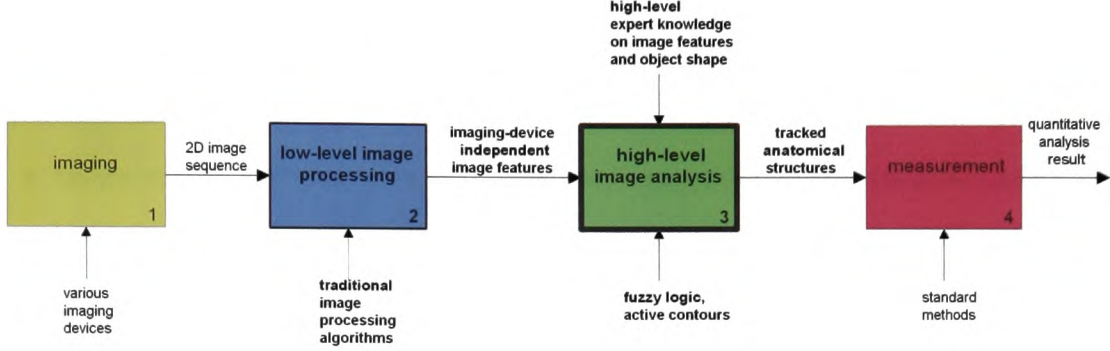
Figure 1.1: Block diagram of the presented system.

The subsequent innovative high-level image analysis is responsible for the actual identification and tracking of an object's contour. A key philosophy of the proposed system is that this task cannot be performed without *a priori* knowledge on both the object itself and its appearance in the image sequence. This primary idea was investigated by HÖWING *et al.* in [1] and developed further by HÖWING *et al.* in [2]. The final component (block 4) performs application specific measurements, usually through generic geometric algorithms, although this is not primary focus of the research.

Another key component of the proposed system is the way uncertainties are handled. The high-level analysis explicitly exploits medical expert knowledge. As this knowledge is often qualitative, vague and may only be acquired in natural-language terms rather than precise parameters, a framework based on fuzzy logic is incorporated in the high-level component (block 3). To perform the actual contour segmentation this novel building block provides a seamless combination of a modified active contour with the fuzzy framework, introducing the original concept of fuzzy active contours, presented by HÖWING *et al.* in [3] and detailed by HÖWING *et al.* in [4].

Specific *a priori* knowledge is excluded from the low-level processing stage to maintain its effectiveness. Furthermore all uncertainties introduced by the image formation

process (block 1) must be propagated to the higher level (block 3). The actual decision on the segmentation of a pixel is deliberately delayed, so as not to bias any decision, based on general assumptions until the higher level, at which point a much more reliable decision can be made which takes specific object knowledge into account.

The main objectives of this research are listed below:

- Investigating and classifying the sources of uncertainties and variation introduced by various imaging devices, deformable and moving objects as well as imprecise object descriptions;

- Investigating the usability of active contour model as a segmentation method within the problem domain;

- Investigating the possibilities of fuzzy logic as a generic framework for both knowledge representation and contour segmentation;

- Establishing a novel theoretical framework combining both active contours and fuzzy logic;

- Considering existing image processing operators to provide image features that can be integrated into the proposed fuzzy active contour;

- Establishing a novel framework for such an integration;

- Development of an innovative knowledge representation that allows for a description of both object and image properties, taking into account the identified uncertainties and variations.

In conclusion, the objectives of the work presented in this thesis recognise the novel philosophies of the previously identified innovative building blocks within the proposed system framework shown in Figure 1.1. Verification of the proposed principles are performed on two complementary applications. The underlying methodology of this verification approach is detailed in the following section.

1.3 Methodology

Computer vision is a wide area were no general purpose solution to even moderately complex problems exists. Many decades of research have demonstrated that it is necessary to define the outline of a vision problem, preferably by formalising a target application. In order to obtain a more general solution an area of similar applications can be defined. If this is envisaged, many characteristics of both images and objects have to be considered carefully.

In this thesis, the area of application considered are firstly surveyed theoretically. These theoretical issues are substantiated by two representative applications presented by HÖWING *et al.* in [2] and [5] respectively.

There is no valid test data or quality assessment criterion for the considered domain, which could allow for a validation of the theoretical properties of the proposed new approach. Hence in a first step, solutions for the example applications are realised using several well-referenced traditional image processing techniques. These solutions demonstrate both the potential and certain limitations of the existing methods.

During the realisation of the example applications, a certain view on image processing systems is adopted which is based on a classification of KASTURI and JAIN and which is introduced in section 2.1.1. It is shown that the realisations of some of the processing stages identified by KASTURI and JAIN are useful. For other stages however, a novel approach is presented, so both the new and the traditional solutions are based on the same system structure. Hence it is possible to compare both traditional and novel approaches when applied to the example applications. This forms the basis for an application-based validation of the new approach and at the same time allows for a demonstration of its benefits over traditional solutions.

1.4 Overview

Chapter 1: Introduces the background and goals of this research.

Chapter 2:   Sets the scope of the project by placing it into a theoretical computer vision context as well as defining an area of application. Two example applications are presented.

Chapter 3:   Summarises the results of literature surveys for the two major scientific areas that are combined in this research: active contours and fuzzy logic. Relevant foundations of both subjects are summarised, making this thesis self-contained.

Chapter 4:   Within this part of the thesis the theoretical analysis of the aforementioned novel approach is given.

Chapter 5:   Validates the theoretical properties of the thesis' main contribution through an experimental analysis of both synthetic and medical image data.

Firstly, many of the major traditional image processing methods are evaluated by elaborating different solutions for the example applications. Subsequently the more recent approach of active contours is adapted and applied. Having analysed the capabilities and deficiencies of existing image processing methods in practice, the novel approach is applied to the example image data.

Chapter 6:   Conclusion and further work.

Appendix A: A brief introduction into the notation of the Structured Analysis and Design Technique.

Appendix B: Selected papers published during this research.

Chapter 2

Quantitative Analysis of Medical Image Sequences

In this chapter the research is positioned theoretically within the diverse area of computer vision. The scope is then set through an application-centred perspective, defining the central terms of the research: *quantitative analysis*, *medical*, and *image sequences*.

The second section provides a brief survey of image and scene properties as relevant to medical image processing. Uncertainty and variation which play an important role in the outlined domain are also investigated.

The third section substantiates what was described in the preceding sections by introducing two complementary applications, which together form a representative basis to verify the validity of the proposed approach.

2.1 Scope of the Project

2.1.1 Computer Vision Context

The ultimate goal of computer vision is the development of an artificial system with scene-interpretation capabilities comparable or even superior to that of humans.

Since the human visual system is not fully understood in detail there is an ongoing discussion on the various aspects of vision, cognition and models of scene structure [6]. Some of the current paradigms are strongly influenced by the more philosophical aspects of perception. An interesting interdisciplinary approach has been proposed by WECHSLER [7] who states that

computational vision

= parallel distributed computation

= parallel distributed representation + parallel distributed processing
+ parallel distributed strategies

= parallel distributed representation + parallel distributed processing
+ active perception

= parallel distributed representation + parallel distributed processing
+ (functional active perception + exploratory active perception)

It is clear and emphasised by the use of the term *computational vision* rather than *computer vision* that this approach cannot be realised with today's technology.

During the development of image processing and computer vision research over the past two decades, several more realistic and hence limited models have been proposed in the literature. There are several possible ways to look at the structure of a computer vision system, that is to describe a system model depending on the purpose of the system, the background and intention of the author.

BALLARD, for example, states that "*Visual perception is the relation of visual input to previously existing* models *of the world.*" [8] He focuses on four categories of representation that are necessary to connect the input to the output.

RADIG puts an emphasis on image *understanding,* by proposing a four-layer structure going from the raw image to a scene interpretation [9]. The model focuses on declarative knowledge and sets aside control.

As with BALLARD and RADIG, the system structure of PINZ possesses several

levels of representation, but in addition processes (algorithms) and control structures are given equivalent importance within the model [10].

From a more practical view point, KASTURI and JAIN give an overview of computer vision [11, 12] and propose the following main subjects of the field:

1. Image formation

2. Segmentation

3. Feature extraction and matching

4. Constraint exploitation and shape recovery

5. Three-dimensional object recognition

6. Dynamic vision

7. Knowledge-based vision

Many text books on computer vision refer to similar subjects [8, 13–21]. This view is adopted in this thesis as it is appropriate for the application area considered and also because it is fully congruent with the methodology described in section 1.3.

2.1.2   Application Area

The following items briefly characterise the area of application within which the contributions of this thesis are validated:

- A *quantitative analysis* of images requires

  - the measurement of anatomically relevant parameters, typically points on or relative to object boundaries, as well as

  - the determination of an anatomically relevant co-ordinate system.

  The alternative is a qualitative analysis where the contents of scenes are described (for example in robot vision) or where whole images are classified (for example in image retrieval systems).

- *Medical images* in this thesis involve

  - known anatomic structures with

  - inter-individual variations in these structures.

  Another important characteristic of medical images is the presence of a high amount of distortions, for example blurry or superimposing structures.

- Imaging *sequences* lead to

  - global motion of whole objects relative to their surroundings, and

  - local object deformation, that is intra-individual variations.

  Imaging sequences in the context of this research are sequences of two-dimensional images taken at discrete time intervals, commonly referred to in the literature as the 2D+t case. Motion and deformation are assumed to only take place in or parallel to the image plane. If this requirement is not upheld then the processing of the third dimension of objects were required, which is beyond the scope of this research as already mentioned in section 1.1

Within this application area, the ultimate goal of an image processing solution is a precise determination of anatomically relevant parameters, in the presence of distortions and variation influences, sources of which are analysed in the next section.

2.2 Sources of Uncertainty and Variation

The problem of relating image structures to the real-world objects they originate from is that there is no unequivocal one-to-one mapping. The image formation process normally involves a loss of information and the introduction of noise. Models on imaging and objects that are used to compensate for these effects are often incomplete or imprecise.

Characteristics or distortions which cannot be addressed through a systematic analysis or their deliberate exclusion from the process introduce a factor of uncertainty to the system. Examples include a vague definition of the objects to be measured as well as unknown parameters of the imaging device.

Natural variations within or between individuals pose another class of problems, which have to be analysed in order to develop a solution that is flexible enough to cover the class variations.

Both uncertainty and variation are particularly relevant, hence this section briefly surveys the main sources of uncertainty resulting from image formation and object models in general. It then delimits the actual image formation characteristics considered in this thesis. Subsequently, the variations introduced by objects in medical imaging sequences, which are especially relevant to this project are surveyed.

2.2.1 Uncertainty in Image Processing

Processing grey-level images involves different kinds of uncertainty at different processing levels. TIZHOOSH identifies uncertainties that stem from the grey-levels themselves, from their relationships and from expert knowledge [22]. For this purpose TIZHOOSH classifies image processing steps similar to those identified by KASTURI and JAIN (cf. section 2.1.1) into three levels:

- low level: image formation and preprocessing

- intermediate level: segmentation, visualisation, description

- high level: analysis, interpretation, recognition

In the literature, intermediate-level methods are not usually described as a separate stage in image processing. Instead they are assigned to either the low or high level, an example of which is the structure proposed by LIEDTKE (cf. Figure 2.1). Here

however, a separation into three levels is performed to identify an additional source of uncertainty.
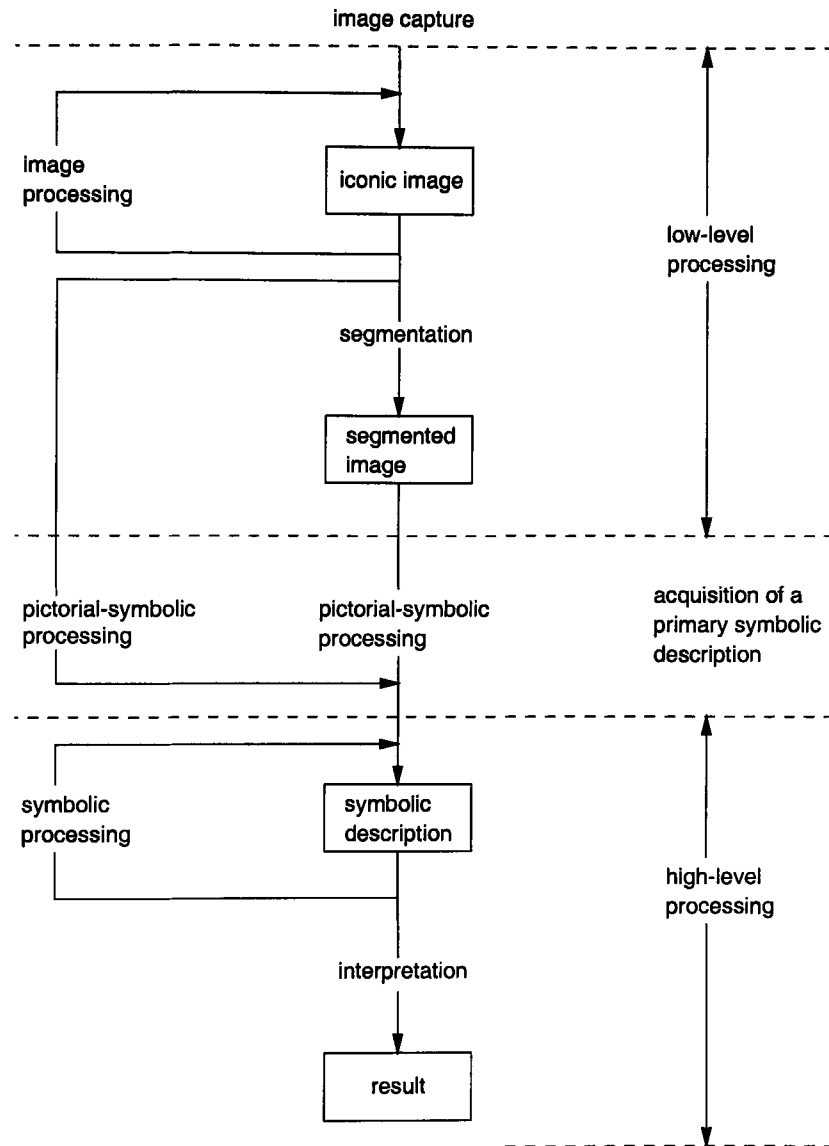
Figure 2.1: Levels of processing in an image analysis system. Adapted from [23].

In explaining the procedure in image understanding, the diagram delimits low-level from high-level processing. It also shows where one may draw the line between the lower iconic and the higher symbolic level of representation.

Greyness Ambiguity

The first source of uncertainty is the image formation process (block 1 in Figure 1.1) as the initial stage of an image processing system. This normally involves a loss of information. An example is the loss of depth information in a 3D to 2D projection. Another example is the blurring of edges during X-ray imaging. This effect can only be compensated for if a precise model of the imaging device is known, which may not always be the case. Similarly the introduction of noise cannot be compensated for completely, leading to further uncertainty at this stage.

Subsequent low-level processing involves the manipulation of the pixels' intensities. Common operations at this level are contrast enhancement, noise reduction, thresholding and most other local operators. To find optimal parameters for these operations is often difficult and usually there is not a unique solution. When determining a threshold value to binarise an image for example, the optimal value depends on the actual histogram as well as on the task that is to be performed.

The intermediate level of processing is concerned with geometrical relations within images and commonly involves local operations. The main concern of intermediate operations is to find object boundaries, trace contours or edges. Questions arise on where the edge or boundary of an object segment actually is.

Indeed, does the object have an edge at all? And if it does, is the edge sharp and well-defined: If the boundary of objects or segments are blurry and ill-defined a crisp decision on segmentation, for example, would result in a loss of information similar to the application of inadequate low-level operations.

Uncertain Knowledge

A priori expert knowledge is normally used at the last, high-level stage of an image processing system. Generally it aids in object recognition, image interpretation, and scene analysis. The basis of these processes is formed by the results of the lower-level

stages. Consequently uncertainty at the high level can occur if uncertainties were not properly handled at the lower-level stages.

In the high-level stage, the expert knowledge itself may be a source of uncertainty. Often an object's model is not known completely resulting in a vagueness of object (class) definitions. A similar uncertainty is introduced by a knowledge representation which is not unique or vague, particularly when knowledge is available in linguistic terms the uncertainty of language leads to uncertain knowledge.

Systematic Uncertainties

The actual implementation of an image processing or computer vision system may introduce further uncertainties. These are usually discretisation problems, hardware or software errors resulting in incorrect or imprecise calculations.

The lack of absolute quantitative measures in image processing is another reason for uncertainties. There exists no absolute measure to assess the image quality after preprocessing for example. This is application dependent and based on subjective assessment of experts or empirical methods.

2.2.2   Image Formation

In general an image is formed when a sensor records a received signal, usually electromagnetic waves, as a two-dimensional function $I(x, y)$, where x and y represent the co-ordinates of the image intensities I. In digital images x, y and I are discrete and a picture element at (x, y) is called a *pixel* (or pel).

Depending on the sensor, the brightness or intensity values I represent for example, the reflectance of light from object surfaces, the temperature of objects or their distance. In this thesis, medical magnetic resonance (MR) and X-ray imaging is considered in particular. Here the image intensities represent structural properties of anatomical objects, that is the distribution of hydrogen and their absorbtion coefficient respec-

tively. X-ray images are characterised by the transparent projection of object densities, while MR images are formed from a local atomic response, allowing slices of objects to be scanned. These particular imaging methods are described in detail in sections 2.3.1 and 2.3.2 respectively. More information on medical imaging in particular can be found in [24].

While a single image is represented by the two-dimensional function $I(x, y)$, image formation may result in more dimensions, obtaining image sequences. These can be devided into

2D	A single two-dimensional image is acquired.
$2\frac{1}{2}$ D	Additional information, interpreting I, is available.
2D+t	A sequence of two-dimensional images is acquired over time.
3D	Depth information may be obtained directly, calculated from the combination of two or more 2D-images or by taking a spatial sequence of two-dimensional image slices.
3D+t	A sequence of 3D-images is acquired over time.

This research project considers the 2D+t case in particular. This means that rather than static scenes dynamic processes involving motion are investigated. What is intuitively regarded as motion may have a number of causes in image sequences. Motion results in a change of image intensities over time. Intensity changes however, may be due to camera motion or object motion, as well as to illumination changes, or changes in object structure, size or shape. As not all of these influences may be considered to be motion, it is important to identify them in an image processing system. In this section changes due to the image formation process are considered, while the next section explores object-induced changes.

As to motion caused by camera (or imaging device in general) or objects the following cases represent the four possibilities for such a dynamic camera/world setup: [11, 12]

1. Stationary camera, stationary objects (SCSO);

2. Stationary camera, moving objects (SCMO);

3. Moving camera, stationary objects (MCSO); and

4. Moving camera, moving objects (MCMO).

The first case involves only static image processing, while the SCMO case allows an analysis of object motion. The MCMO case is the most challenging and usually requires additional information about the motion of either the camera or the object to fulfil computer vision tasks. This research project primarily considers the SCMO case which is most common in medical applications, where motion is involved.

2.2.3 Object Characteristics

Regarding objects within image sequences the following variations are possible:

- **Intra-individual** Variations occur within the same individual over the course of an image sequence (over time and/or space that is).

- **Inter-individual** Variations of the same anatomical object that occur over different individuals.

- **Motion** Global change in position of an object, relative to the image co-ordinate system. Motion can be divided into translation and rotation.

- **Shape variation** Local changes in the 2D contour of an object lead to deformation.

It is important to emphasise that the terms used relate to the *appearance of a two-dimensional object boundary* in an image sequence, which may lead to a different classification than a look at the physical object in three dimensions would give. In the real world a bone for example, would be classified as a rigid object, because it cannot be deformed. In a sliced MR image sequence however, the 2D contour of that bone may

undergo deformations from one slice to another as the position of the slice changes (see section 2.3.1). Similarly in an X-ray image sequence the projection of a rotating bone may lead to a deformation of its outline, when the rotation is not restricted to the axis perpendicular to the view plane.

Tables 2.1 to 2.3 survey a number of object classes, derived from the possible combinations of the above variations. For each class a medical example is provided.

Object class	Object characteristics						Medical example
	Intra-individual				Inter-individual		
	Motion		Shape variation		Shape variation		
	Translation	Rotation	Rigid	Deformable	Rigid	Deformable	
artificial reference structure	o	o	×	–	×	–	fiducial marker

Table 2.1: Rigid objects. Definitions of object variations and resulting object class. Legend: × Yes, – No, o Yes or No

Object class	Object characteristics						Medical example
	Intra-individual				Inter-individual		
	Motion		Shape variation		Shape variation		
	Translation	Rotation	Rigid	Deformable	Rigid	Deformable	
fixed anatomical reference structure	–	–	×	–	–	×	bone, fixed relative to imaging device
anatomical reference structure, moving in the view plane	–	×	×	–	–	×	bone, restricted to move only parallel to view plane
	×	–	×	–	–	×	
	×	×	×	–	–	×	
fixed object (rigid in 3D)	–	–	×	–	–	×	bone, fixed relative to reference structure
object (rigid in 3D), moving in the view plane	–	×	×	–	–	×	bone, restricted to move only parallel to view plane
	×	–	×	–	–	×	
	×	×	×	–	–	×	

Table 2.2: Inter-individually deformable objects. Definitions of object variations and resulting object classes. Legend: × Yes, – No

Object class	Object characteristics						Medical example
	Intra-individual				Inter-individual		
	Motion		Shape variation		Shape variation		
	Translation	Rotation	Rigid	Deformable	Rigid	Deformable	
object (rigid in 3D), rotating around x- and/or y-axis	–	×	–	×	–	×	bone, moving out of the view plane
	×	–	–	×	–	×	
	×	×	–	×	–	×	
soft objects, connected to reference structure	–	–	–	×	–	×	tissue, attached to reference bone
soft object, not connected to reference structure	–	×	–	×	–	×	
	×	–	–	×	–	×	
	×	×	–	×	–	×	

Table 2.3: Intra-individually deformable objects. Definitions of object variations and resulting object classes. Legend: × Yes, – No

2.3   Example Applications

This section substantiates the discussions of the preceding sections by introducing two applications from different imaging as well as application domains. Together the applications address all of the subjects mentioned above and are regarded as a representative basis to verify the validity of the proposed approach.

2.3.1   Analysis of Carpal Bones in Magnetic Resonance Imaging Sequences

2.3.1.1   Clinical Motivation

To date the diagnosis of ligament lesions of carpal bones relies on a qualitative examination of the patient's wrist. This section presents a novel system where sequences of magnetic resonance images are automatically analysed to measure the motion of seven wrist bones. Resulting motion graphs provide a quantitative basis for diagnostic as

well as scientific purposes. As the imaging method is non-invasive up to twelve wrist positions can be measured giving a detailed insight into the bone's motion.

In many cases articular damage cannot be diagnosed through an examination of a single image. A motion analysis of a joint's bones might be necessary to make a reliable diagnosis [25,26]. Examples are lesions of the ligaments and cartilage of the knee or in the cervical and lumbar regions of the vertebra. The developed system enables the diagnosis of lesions of the ligaments of the wrist (carpal instabilities [27]), and is particularly well-suited to aid in the diagnosis of the scapho-lunate instability. This damage is a common injury after accidents involving the wrist. The lesion occurs when the ligaments between the Scaphoid and the Lunate are torn [28].

Availability of Nuclear Magnetic Resonance Imaging allows scanning of entire sequences of images of bones and joints without harmful dosage of radiation. Analysis of such sequences allow a much more reliable diagnosis of lesions of the ligaments compared to methods in use today, such as single X-ray images [25–27]. However, a necessary scanning procedure with a sufficient number of positions requires approximately 100 2D images for every patient. The manual evaluation of such a large number of images in the daily medical diagnostic is not feasible.

The aim of the developed solution is the automatic processing of these images in order to obtain motion graphs which allow an easier medical diagnosis. For the recognition of lesions of the ligaments, the representation of translation and rotation of the carpal bones with respect to a co-ordinate system defined by the radius proved to be most suitable. Using cadaveric specimens such investigations have been carried out by implantation of markers [28]. Because of the considerable exposure to radiation, in vivo analysis of such movements has been carried out only with very coarse resolution. The method proposed here allows for a much finer resolution of the bone movement.

For the measurement of translation and rotation for each bone its major axis and centroid is determined. The measurement is performed relative to an anatomic co-ordinate system defined by the distal end of the Radius bone (cf. Figure 2.2).
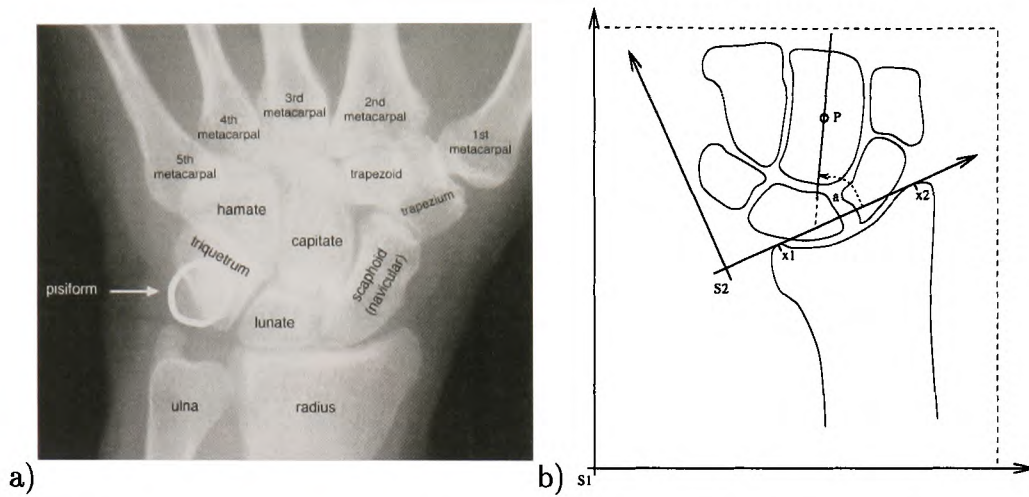
Figure 2.2: a) X-ray image of the wrist, identifying the bones. b) Parameter definition: P: centroid position, a: angle of rotation, S_1: image co-ordinate system, S_2: target co-ordinate system, x_1 and x_2: reference points (landmarks) defining S_2

2.3.1.2 Image Formation

There are many system parameters in an MRI device [24]. For automated image processing, it eases the design of a solution considerably if image formation parameters do not change. Hence together with the MR operators, useful settings were determined and fixed, guaranteeing a constant appearance of bones and surrounding tissue. Examples are shown in Figure 2.3.
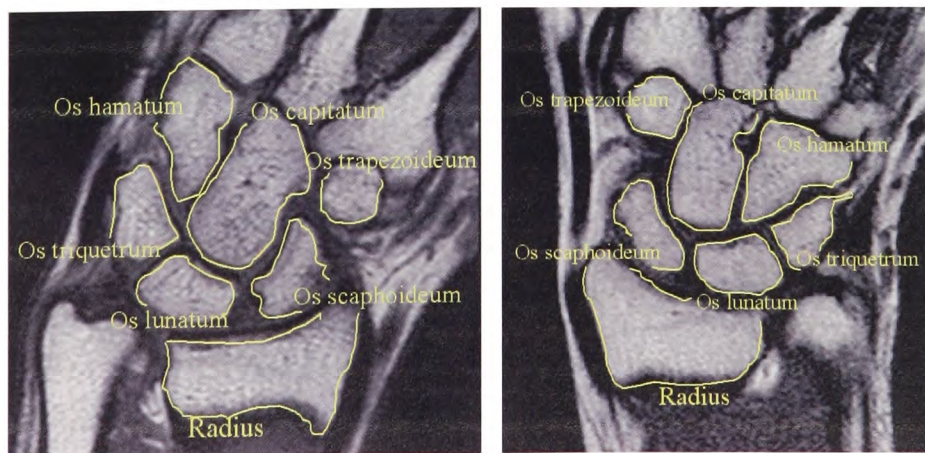
Figure 2.3: MR images of the left and right hand. Relevant bones are identified.

Apart from the imaging device, the test person's hand also has to be fixed to guarantee reproducible results. A restriction of the movement to a plane is achieved by fixing the test person's lower arm and hand in a device especially developed for this examination, as shown in Figure 2.4.
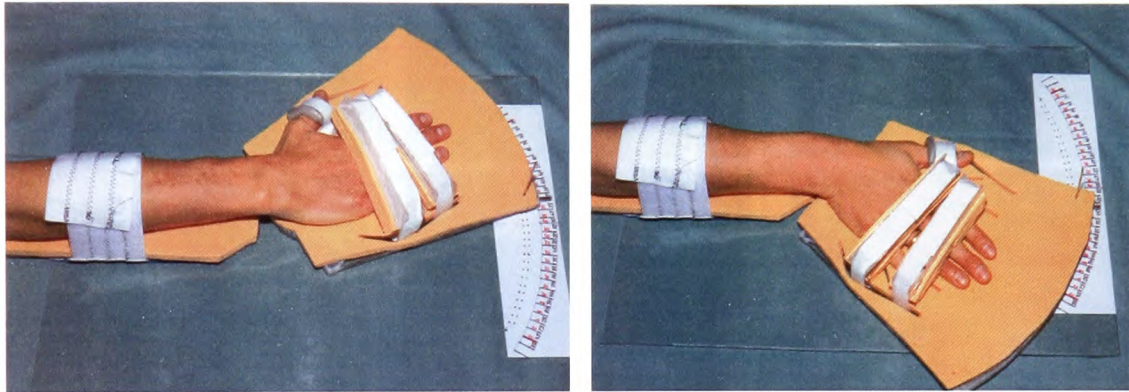
Figure 2.4: Fixation device to restrict the movement of the hand to a plane.

Depending on the flexibility of each patient up to 10 different positions of the wrist will be scanned. For each wrist position 12 layers of the hand are acquired (cf. Figure 2.5 and 2.6). The images are T1-weighted, the distance between layers is 1.5 mm. The resulting grey-level images consist of 512 by 512 pixels, with a depth of 10 bit.

2.3.1.3 Object Characteristics

Compared to other applications involving for example the analysis of the knee or spine, a motion analysis of wrist bones is more difficult because there are many bones with a similar shape which complicates their identification. Furthermore, some of the bones may tilt, that is they may rotate about an axis that is not perpendicular to the viewing plane. This results in a variable appearance of the bones in the sliced magnetic resonance images. The variations and the resulting object classes are summarised in Table 2.4.
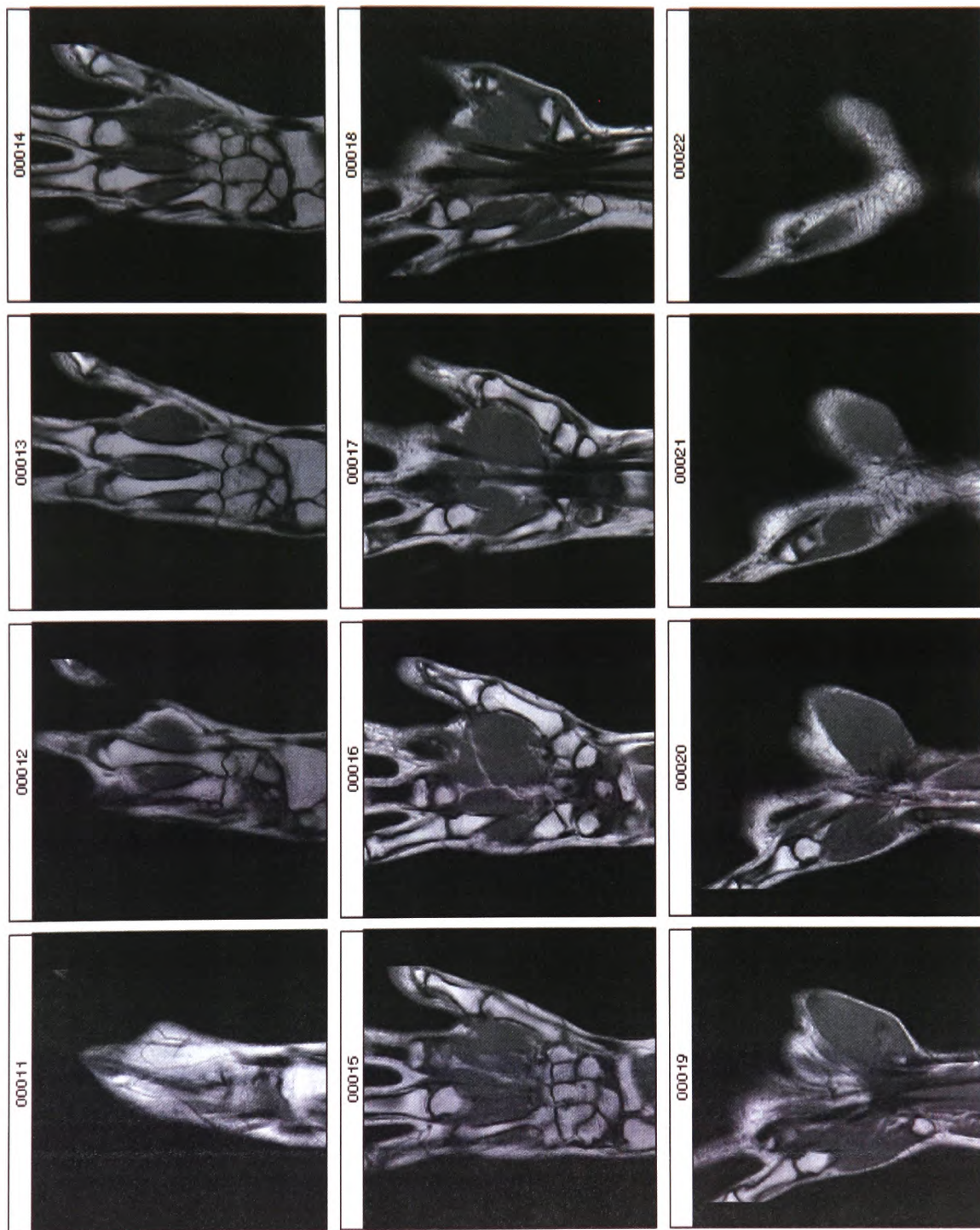
Figure 2.5: All slices of one hand in the neutral (straight) position.
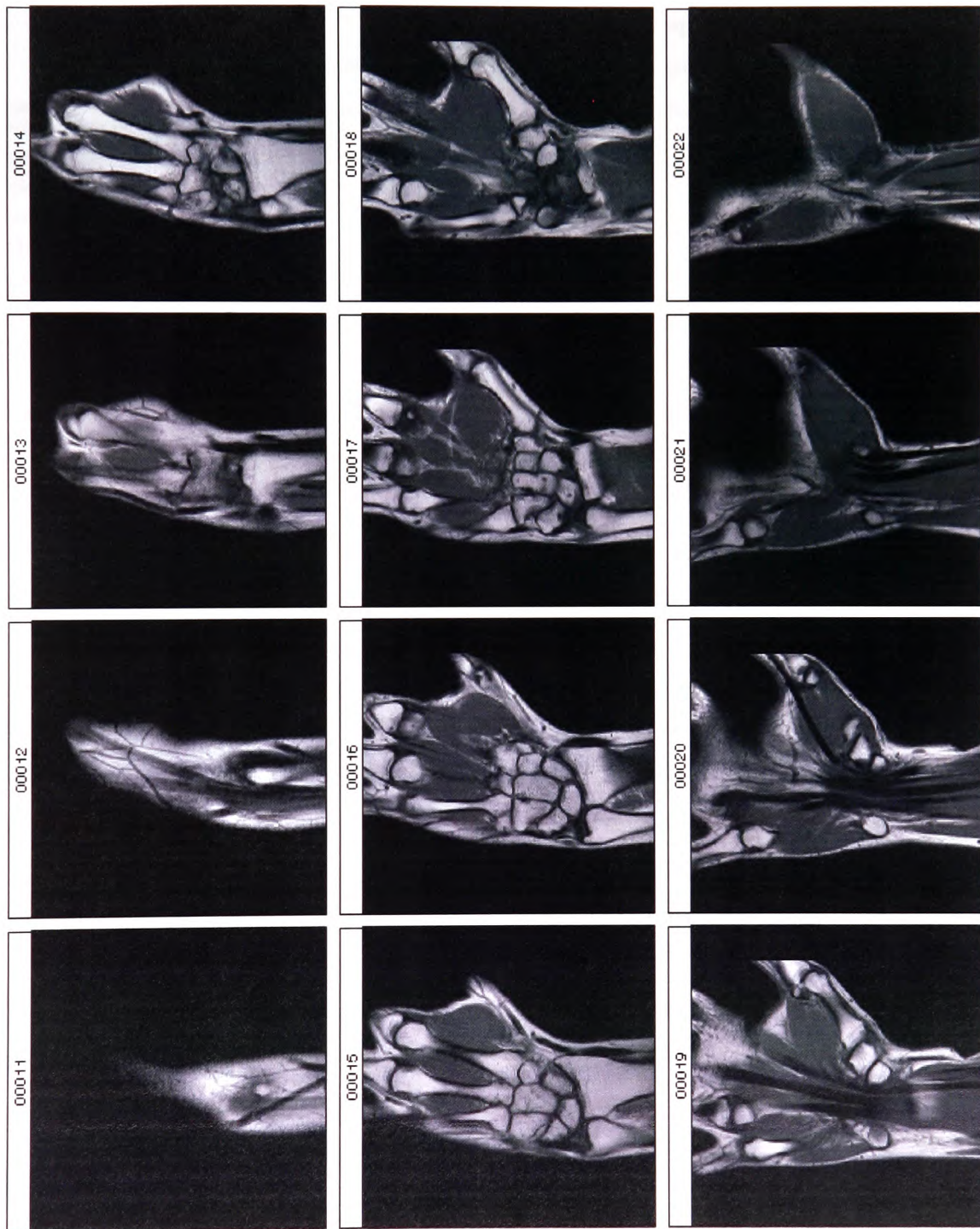
Figure 2.6: All slices of one hand in a radial (inward) position.

Object class	Object characteristics						Wrist bones
	Intra-individual				Inter-individual Shape variation		
	Motion		Shape variation				
	Trans-lation	Rotation	Rigid	Deform-able	Rigid	Deform-able	
anatomical reference structure, moving in the view plane	×	×	×	–	–	×	Radius bone
object (rigid in 3D), moving in the view plane	×	×	–	×	–	×	carpal bones

Table 2.4: Variations of the wrist bones. Legend: × Yes, – No

2.3.2 Analysis of the Vocal Tract in X-ray Imaging Sequences

2.3.2.1 Scientific Motivation

Articulatory phonetics is a branch of linguistic science that is concerned with the very complex dynamic characteristics of the organs of the human vocal tract (cf. Figure 2.7).

> '*The probability that an articulator will move parallel or anti-parallel to the preferred direction during speech sound production is dependent on the movements' orientation of the production of his neighbouring sound which in turn is influenced by the production of the neighbouring sound and so on. This chain of effects develops because the vocal tract is a dynamic network system in which different articulatory parts effect each other through interaction.*'[29]

Understanding this complex motion and interrelation of articulatory organs is an important basis for understanding human speech production. Apart from being a contribution to basic linguistic research, this knowledge is valuable for example to speech

therapy and speech recognition [30–32].

To analyse the articulatory organs it is necessary to determine the position of 12 characteristic parameters of the vocal tract defined on the midsagittal plane (cf. Figure 2.8). For this purpose X-ray image sequences have been taken from different speakers while uttering certain (arabic) syllables. Figure 2.10 shows two examples.
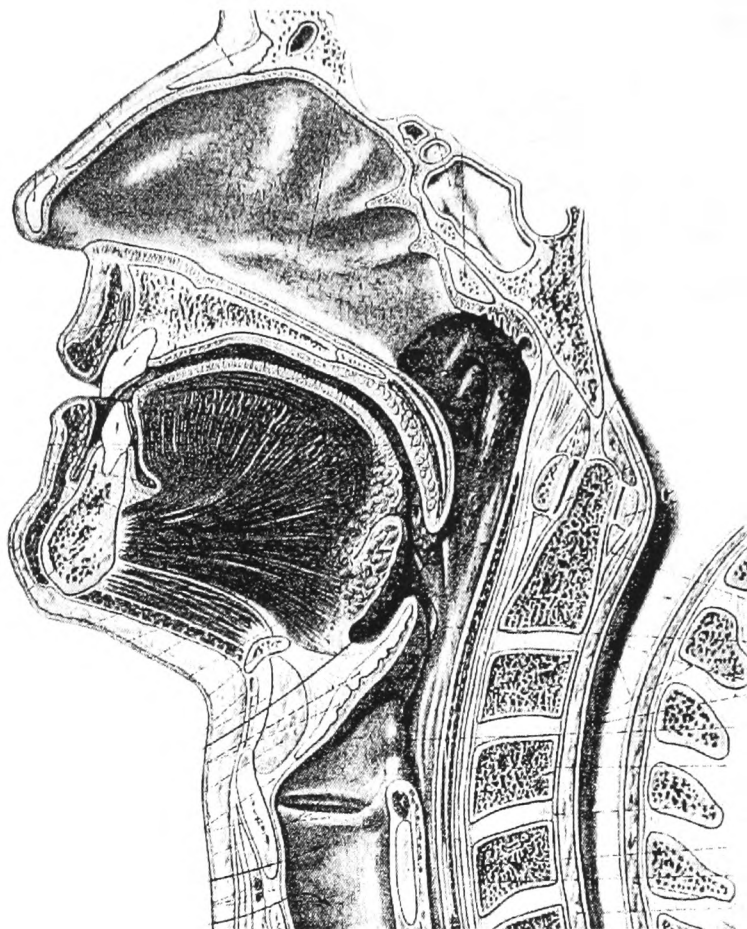
Figure 2.7: Anatomy of the vocal tract. From [33]

For a statistically reliable analysis a very large number of images have to be processed. A manual measurement would be prohibitively time-consuming. Furthermore, an automated processing eliminates inter- and intra-observer variabilities from the measuring process.

1	Lip positions
2	Lip opening
3	Front teeth
4	Tip of tongue
5	Front tongue
6	Middle tongue
7	Back tongue
8	Velum
9	Epiglottis
10	Hyoid bone
11	Glottal narrowing
12	Glottis
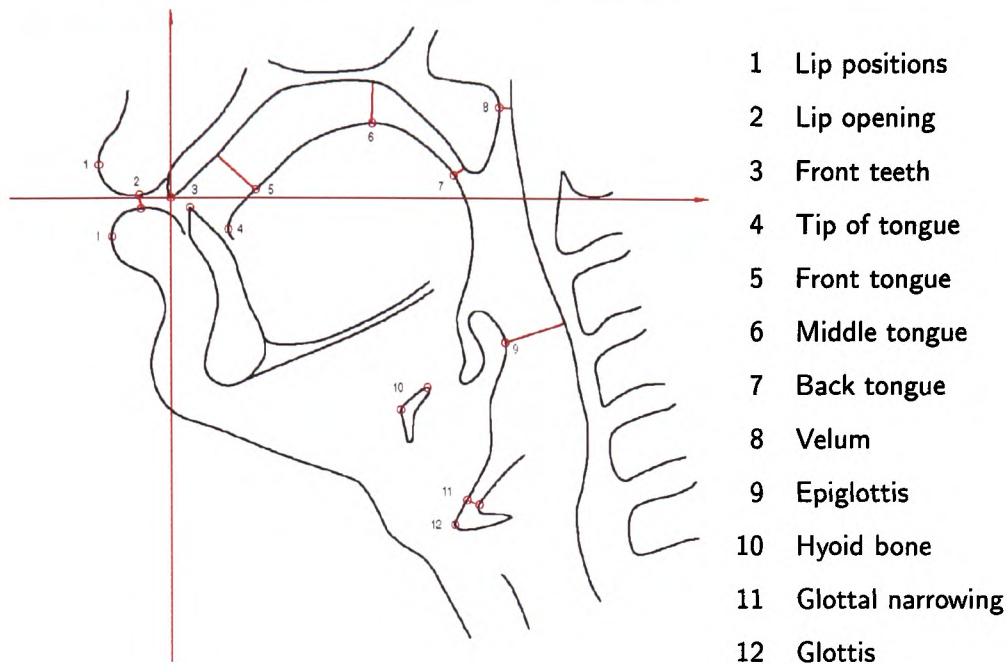
Figure 2.8: Characteristic parameters of the vocal tract.

2.3.2.2   Image Formation

With the MR imaging-based application presented in section 2.3.1 it was possible to tailor the imaging conditions in order to ease the subsequent image processing. The X-ray image sequences of the vocal tract however, were taken in the early eighties in France and supplied to the author by linguists. Initially the sequences existed only on 35mm celluloid films. Neither the precise set-up of the experiments nor details about the imaging device is known. Nowadays it would be ethically unacceptable to expose test persons to radiation solely for speech scientific purposes. It is therefore not possible to create new image sequences with known parameters.

At the time the image sequences were taken it was already envisaged to analyse the movements of the articulators. The upper part of the head was fixed, but this fixation was not perfect. In some sequences the test person's head moves. Consequently intensities of the resulting images vary considerably in those sections.

To facilitate measurement small leaden markers were attached to the lips (upper lip, lower lip and in the corner of the mouth) of the test persons. These markers are

visible as dark circles. Unfortunately the markers were not always fixed properly so they might change position or visible shape and so hinder the image processing more than they are of use.

The particular X-ray imaging leads to further challenging images properties. Generally all edges are blurry and have a low local contrast. To make soft tissue visible a contrast agent was given orally. Salivation however, causes an inhomogeneous and constantly changing distribution of the contrast agent. This in turn affects the appearance of soft objects such as the tongue. Edges and textures of such objects are therefore not constant.

Digital image processing required the original 35mm films to be scanned and digitised. In addition to inherent visual artefacts such as scratches and dirt on the films, this largely analog process turned out to be another source of distortion. The signal dynamic range was low resulting in a further reduced contrast and a relatively high signal noise. As a high-quality digitisation was not available at the time, the celluloid was first scanned using a professional film-to-video scanner (cf. Figure 2.9a).

a)   b)
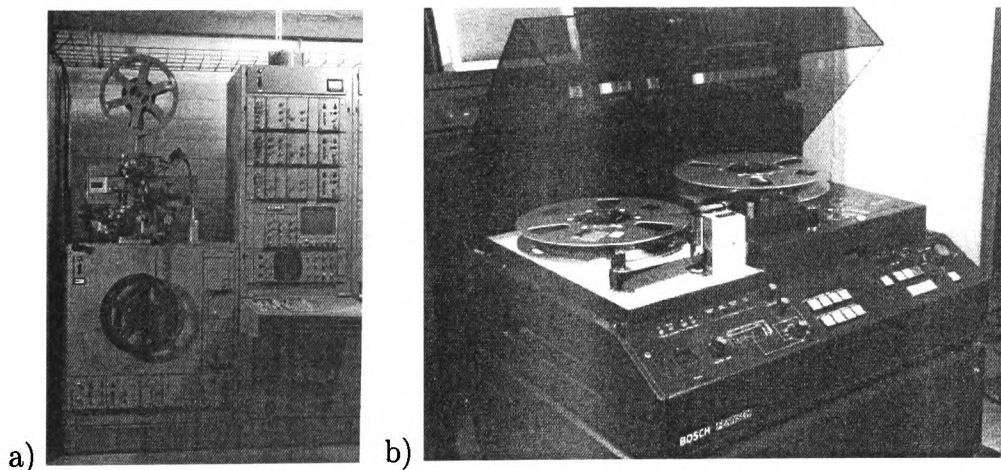
Figure 2.9: a) 35mm film-to-video scanner, b) 2-inch magnetic tape recorder.

The analog video signal was stored on 2-inch magnetic tapes (cf. Figure 2.9b) from which it was subsequently digitised in full CCITT PAL resolution (720x568 pixels). Two examples of the resulting digital images are shown in Figure 2.10.
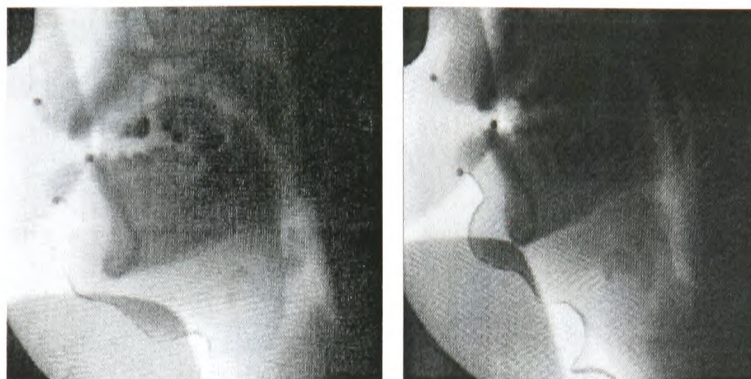
Figure 2.10: Digitised X-ray images of the vocal tract from two different test persons.

2.3.2.3 Object Characteristics

The objects in this application are defined by the parameter points that have to be measured rather than solely through anatomical knowledge. In theory some of the points could be detected directly, that is without identifying the object boundaries first. This approach depends on the particular image processing method and section 5.2.2.1 will evaluate such methods.

Other points however are defined by the minimal Euclidean distance between opposite sections of two articulators, rather than by a single characteristic point on the object (cf. Table 2.5 and Figure 2.11). Consequently to measure the above mentioned 12 parameters the precise location of 22 relevant boundary points have to be determined.

	Parameter	*Articulator(-sections) involved*
2	Lip opening	Upper and lower lips
5	Front tongue	Front tongue and alveolus
6	Middle tongue	Middle tongue and palate
7	Back tongue	Back tongue and velum
8	Velum	Velum and pharynx wall
9	Epiglottis	Epiglottis and pharynx wall
11	Glottal narrowing	Glottal walls

Table 2.5: Boundary points in the vocal tract defined by a minimal distance.

A set of midsaggital sketches in Figure 2.12 for arabic consonants provide an exam-

ple for the variability of the articulators, in particular of the lips, the tongue and the velum. A complete classification of the vocal tract's objects is provided in Table 2.6.
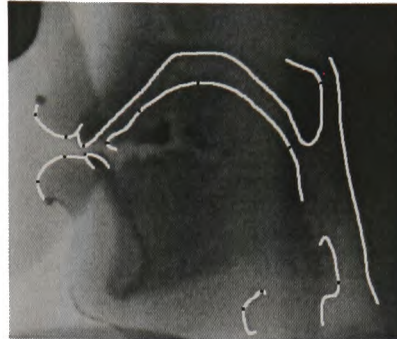
Figure 2.11: Complete object boundaries as necessary to determine the measurement points.
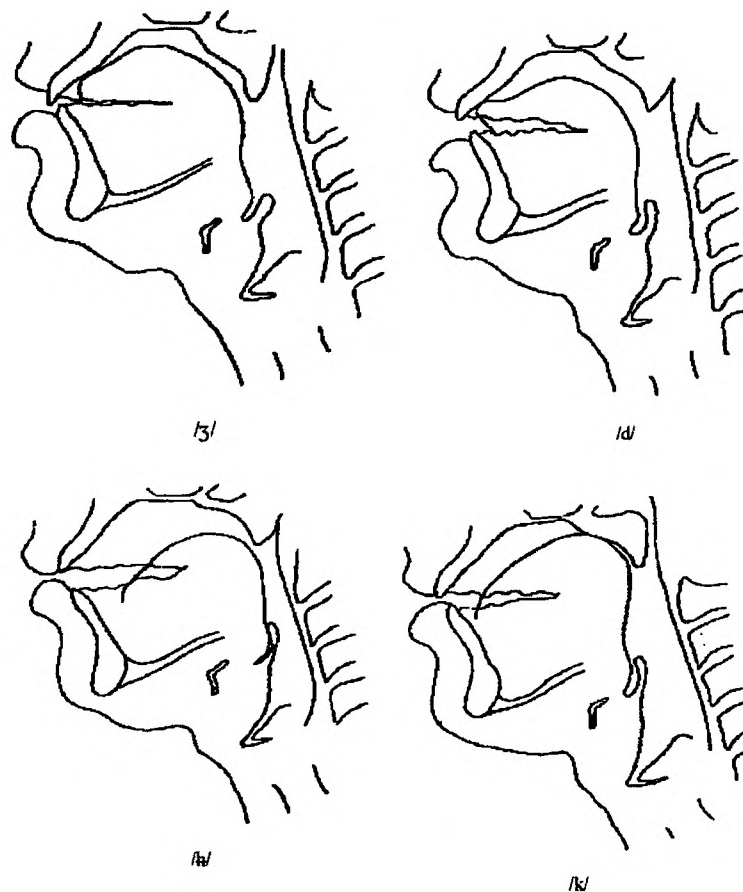
Figure 2.12: Sketches of the vocal tract for some arabic consonants.

Object class	Object characteristics						Articulators
	Intra-individual Motion		Shape variation		Inter-individual Shape variation		
	Trans-lation	Rotation	Rigid	Deform-able	Rigid	Deform-able	
artificial reference structure	×	×	×	–	×	–	lip marker
fixed anatomical reference structure	–	–	×	–	–	×	upper front tooth
fixed object (rigid in 3D)	–	–	×	–	–	×	alveolus, palate, pharynx
object (rigid in 3D), moving in the view plane	×	×	×	–	–	×	lower front tooth, hyoid bone
soft objects, connected to reference structure	–	–	–	×	–	×	velum, upper lip, epiglottis
soft object, not connected to reference structure	×	×	–	×	–	×	lower lip, tongue, glottis

Table 2.6: Variations of objects in the vocal tract. Legend: × Yes, – No

2.3.3   Summary of Example Applications

The preceding sections presented two example applications that lie within the scope of this work as detailed in section 2.1.

The clinical analysis of MR image sequences and the scentific measurement of X-ray sequences are complementary with respect to the variation classes identified in section 2.2.2 (image formation) and section 2.2.3 (object characteristics).

While MRI produces image slices with good contrast and edge quality, analog X-ray imaging results in low-contrasted blurry edges, with superimposing structures.

The objects to be analysed are both rigid and deformable. While the MR application focuses on inter-individual deformation (carpal bones), the X-ray analysis is mainly concerned with intra-individual deformation of, for example, the tongue.

Both applications involve the challenging properties investigated in section 2.2 and provide a basis for both the evaluation of traditional image processing methods and the validation of the novel approach proposed in this work.

Chapter 3

Related Activities and Literature Surveys

While the analysis of medical images and image sequences dates back to the early 1970's, it is still an active field of research and has led to various medical applications, the most prominent being the analysis of the heart. In particular volumetric and dynamic analysis of the left ventricle has led to numerous publications over the last three decades [34–38]. A similarly relevant application is the assessment of the state of coronary arteries [39–42], whilst among the first anatomic structures to be processed digitally were the lungs [8, 43]. Other applications have been concerned with rigid structures of the head [44] and the non-rigid tongue [45, 46].

In all these medical image analysis applications, the main task has been the segmentation of anatomically relevant structures [47–55]. There currently exists no single segmentation method that yields acceptable results for every type of medical image material and application. Multiple techniques are frequently combined to improve the results. Methods do exist that can be applied to a variety of data, however, those that are specialized to particular applications achieve better performance by taking into account *a priori* knowledge. Some of these existing methods are referenced in this section.

Existing approaches can be divided into two categories. The first group contains pixel classification methods such as thresholding, region growing, classifiers and clustering algorithms. The second group includes methods based on deformation such as atlas warping as well as a number of important deformable models.

Early approaches were very much application specific and low-level oriented [34, 35, 39, 45]. Typically methods such as thresholding and region growing were used together with other local operators. For medical segmentation variations on classical thresholding have recently been proposed that incorporate information based on local intensities [56] and connectivity [57]. Region growing is mainly used for detection of small simple structures such as tumors and lesions [58].

Supervised classifiers require a manually segmented training set which is used for automatically segmenting new data. Here statistical methods such as the maximum likelihood classifier play an important role. Examples for improving and applying such classifiers for medical image analysis include [59] and [60]. Also artificial neural networks (ANNs) can be applied as supervised classifiers [61].

When no training data is available, unsupervised clustering methods may be used for pixel classification. Markov random fields (MRF) are often incorporated into clustering algorithms such as the k-means algorithm under a Bayesian prior model [62, 63]. MRF modeling itself is not a segmentation method but a statistical model, with segmentation obtained by maximising the *a posteriori* probability of the segmentation given the image data using iterative methods [64]. Other non-statistic clustering methods such as unsupervised ANNs and fuzzy clustering [65], are also available.

In contrast to pixel classification methods, atlas-guided approaches treat segmentation as a registration problem. Here a pre-segmented template image is warped to a target image using either a linear or non-linear transformation. This method has been applied mainly to MR brain images [66], though accurate segmentation of complex structures is difficult due to anatomical variations. Using probabilistic atlases can help to model these variations [67].

It was recognised that models could improve segmentation results [50], particularly with often noisy medical images [51]. The exploitation of *a priori* knowledge includes specific models concerning the image formation process of the object to be analysed [36, 40, 41, 43, 68], and also expert systems were developed to provide explicit knowledge representation schemes [42, 69].

More general approaches which attempt to model boundary or shape properties [34, 70, 71] have provided a rich research area in medical image analysis. The representation and detection of deformable anatomic structures has recently attracted much research interest [38, 72, 73]. An extensive analysis of current research in this area is given in [53], and other relevant contributions may be found in [52, 54, 55].

An influential deformable boundary model, the so-called active contour was proposed by KASS et al. [74] (cf. section 3.1). Later approaches include deformable templates where size and relationships of object subparts are represented in parametrized templates [38, 75–77].

STAIB and DUNCAN [78, 79] augmented the boundary finding process with *a priori* probability information representing the mean shape and natural variation of the object to be segmented. Fourier descriptors were used as model parameters to represent boundaries. SZÉKELY et al. [80] also developed Fourier parametrized models.

COOTES et al. [81] combined deformable shape descriptors with statistical modal analysis. Their active shape model obtains characteristic shape variations from a training set of boundary points. An example medical application published in [82] trains active shapes with the outlines of finger bones in single X-ray images for age assessment.

The novel approach presented in this thesis is partially based on principles of active contour models, which are discussed in depth in section 3.1. Other important aspects of this thesis are related to representing and processing uncertainties. Many methods have been developed in this field [83], the most relevant of them being Bayesian probability theory and fuzzy set theory.

In Bayesian statistics all uncertainties are represented by probabilities. Probability

is regarded as a subjective measure, representing a degree of belief that something will occur. Probabilistic methods play an important role in medical image segmentation, where stochastic uncertainties are handled. Soft segmentations for example can be obtained in statistics through the use of probability functions [60, 84]. Particularly in cases where training data is available probabilistic methods have been applied successfully. Some important approaches were developed in the areas of classifiers, clustering and deformable models, as mentioned earlier in this section. In [78–81] uncertain knowledge of shape is included by incorporating prior probability distributions obtained from a training set. Other relevant references include [59, 85–94]. For a comprehensive survey of the field consider [53, 95, 96].

In this thesis an explicit, linguistic representation is pursued, while training data is considered unavailable. In these circumstances fuzzy logic is adequate and will be discussed fully in section 3.2.

3.1   Survey of Active Contour Models

Active contour models were first introduced by KASS, WITKIN and TERZOPOULOS [74], and are often referred to as the classic *snake* or *deformable contour model*. Active contours are energy-minimising splines or polygons guided by internal and external forces that pull them towards image features during an optimisation process. They dynamically segment an object contour by locking onto nearby edges and localising them accurately.

Applications of active contour models include line and edge detection, detection of subjective contours, motion tracking, stereo matching, and interactive interpretation of image scenes with user-imposed constraints, in the area of computer vision, computer graphics, and computer-aided geometric design [52, 97–101]. More recently active contour models were applied in computer-assisted medical image analysis [46, 72, 102–107].

3.1.1 General Concept

The traditional model is based on a spline with controlled continuity, providing piece-wise smoothness constraints as internal spline forces and thus regularising the deformation of the model in terms of its elasticity and bending. The representation of the traditional active contour model however, is not spline-based during the deformation process, but only for the final interpolation of the result. The external image forces push the model towards salient features such as lines or edges. External constraint forces are responsible for pulling the model near a desired local energy minimum using appropriate automatic attentional mechanisms, or high-level interpretation. In the absence of such mechanisms, interactive approaches like the *snake pit* [74] can be used, providing an interactive approach for defining pushing and pulling forces in the image scene via *spring* and *volcano* forces.

The key point of active contour models is the design and optimisation of suitable energy functions whose local minima comprise a set of alternative solutions which can be based on *a priori* knowledge of the approximate shape, size, location, and motion of the object under investigation, and on a user-defined initial estimate of the object's contour.

During an optimisation process which was originally formulated within a *Euler-Lagrangian* setting for the traditional model, the internal and external image and constraint forces are adjusted to find the desired local optimum causing a suitable deformation of the active contour model.

3.1.1.1 Continuous Spline Representation

An active contour is based on a parameterised contour $\mathbf{v}(s), s \in [0,1]$. Closed contours are obtained by making the contour periodic, for example by setting $\mathbf{v}(0) = \mathbf{v}(1)$. An energy function ε_{snake} is formulated to obtain an estimate of the quality of the model in terms of its internal forces ε_{int} and external forces, such as underlying images forces ε_{image} and user-constrained forces ε_{con}. The energy function integrates a weighted linear

combination of the internal and external forces over the spline contour:

$$\mathcal{E}_{snake} = \int_0^1 w_{int}\mathcal{E}_{int}(\mathbf{v}(s)) + w_{image}\mathcal{E}_{image}(\mathbf{v}(s)) + w_{con}\mathcal{E}_{con}(\mathbf{v}(s)) \, ds \qquad (3.1)$$

As there is no consistency in the literature as to what kind of forces are to be designated as internal, external or constraint forces, the following definition is used throughout this thesis. The snake's overall energy function comprises an internal as well as an external component:

$$\mathcal{E}_{snake} = \int_0^1 w_{int}\mathcal{E}_{int}(\mathbf{v}(s)) + w_{ext}\mathcal{E}_{ext}(\mathbf{v}(s)) \, ds \qquad (3.2)$$

Internal forces are formulated through $\mathcal{E}_{int}(\mathbf{v}(s))$ which represents all general characteristics that an arbitrary object might have, such as smoothness. Image features as well as object-specific constraints are regarded as being external to the spline and are hence grouped into $\mathcal{E}_{ext}(\mathbf{v}(s))$:

$$\mathcal{E}_{ext}(\mathbf{v}(s)) = w_{image}\mathcal{E}_{image}(\mathbf{v}(s)) + w_{con}\mathcal{E}_{con}(\mathbf{v}(s)) \qquad (3.3)$$

This is conform with the view that an energy function can also be regarded as the compromise between internal and external contour shape quality. In this compromise the weighting parameters w_{int}, w_{ext}, w_{image}, and w_{con} control the relative influence of the energy components and are generally determined by a process of trial-and-error.

Moving the vertices of an active contour leads to a change in energy, which transforms the segmentation problem into an optimisation problem.

3.1.1.2 Discrete Polygonal Representation

The continuous energy function $\mathcal{E}_{snake}$ is usually discretised by replacing the integral by summation, leading to a discrete energy function E_{snake}. In the following a discrete formulation of the active contour model is used, which is based on a polygonal

representation of a contour, as shown in Figure 3.1.

For each image, the algorithm requires an initial polygon $\mathbf{P} = (\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_{N-1})$ consisting of N vertices $\mathbf{p}_i = (x_i, y_i)$, where x_i and y_i are the spatial co-ordinates of $\mathbf{p}_i$. The detected boundary is represented by the polygon $\mathbf{Q} = (\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_{N-1})$, with $\mathbf{q}_i = (x_i, y_i)$. Each $\mathbf{q}_i$ is selected from a set of candidates $\mathbf{C}_i = (\mathbf{c}_{i,0}, \mathbf{c}_{i,1}, \ldots, \mathbf{c}_{i,M-1})$. In many applications the candidates $\mathbf{c}_{i,j} = (x_j, y_j)$ are uniformly sampled along a search line normal to the initial polygon and intersecting $\mathbf{p}_i$. Tracking of the contour is achieved by processing a sequence frame by frame and taking the resulting $\mathbf{Q}_t$ as the initial estimated contour $\mathbf{P}_{t+1}$ for the next frame. A closed contour is obtained by setting $\mathbf{q}_0 = \mathbf{q}_N$ and hence $\mathbf{p}_0 = \mathbf{p}_N$ and $\mathbf{C}_0 = \mathbf{C}_N$.
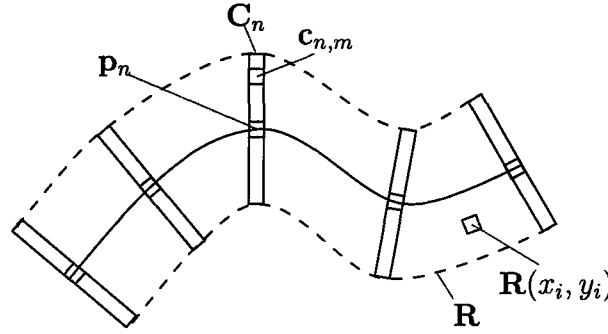
Figure 3.1: Polygonal contour representation.

From this polygonal representation, it is possible to formulate an appropriate discrete energy function E_{snake} for the object, which can subsequently be minimised in order to obtain the desired contour $\mathbf{Q}$. This minimisation is achieved by selecting an optimal set of vertices from the candidates which have been sampled in a region $\mathbf{R}$, around an initial contour.

Analogous to eq. (3.2) E_{snake} is decomposed into two components:

$$E_{snake}(\mathbf{Q}) = \sum_{n=0}^{N-1} w_{int} E_{int}(\mathbf{q}_n) + w_{ext} E_{ext}(\mathbf{q}_n) \tag{3.4}$$

with the external energy being

$$E_{ext}(\mathbf{q}_n) = w_{image}E_{image}(\mathbf{q}_n) + w_{con}E_{con}(\mathbf{q}_n). \qquad (3.5)$$

where E_{image} and E_{con} are the discrete equivalents of the analogue components ε_{image} and ε_{con}.

3.1.2 Modelling Object Characteristics

The classic formulations as well as some modifications and extensions of internal and external active contour forces used to model object characteristics will be summarised in the following.

3.1.2.1 General Characteristics (Internal Energy Terms)

The classical active contour models general elastic deformation as well as bending characteristics through Eq. (3.6) and (3.7) respectively.

$$\varepsilon_{elast}(\mathbf{v}(s)) = \|\mathbf{v}'(s)\|^2 \qquad (3.6)$$

$$\varepsilon_{bend}(\mathbf{v}(s)) = \|\mathbf{v}''(s)\|^2 \qquad (3.7)$$

Here the elastic energy ε_{elast} of the contour is modelled by the first order derivative term $\mathbf{v}'(s)$ which makes the snake behave like a membrane. The second order derivative term $\mathbf{v}''(s)$ represents the contour's bending energy ε_{bend}, leading to a thin-plate like behaviour of the active contour.

Both properties are integrated in the internal energy ε_{int} through a weighted sum:

$$\varepsilon_{int}(\mathbf{v}(s)) = w_{elast}\varepsilon_{elast}(\mathbf{v}(s)) + w_{bend}\varepsilon_{bend}(\mathbf{v}(s)) \qquad (3.8)$$

With this model an object's general characteristics can be parameterised through the weights w_{elast} and w_{bend}. Increasing w_{elast} forces contour vertices to move closer

together, while decreasing the parameter allows the contour to develop gaps. The smoothness of an object is parameterised through w_{bend} where higher values decrease the model's flexibility resulting in a contour with less corners and vice versa.

Discrete approximations of Eq. (3.6) and (3.7) were suggested by [74, 108]:

$$E_{elast}(\mathbf{q}_n) = \|\mathbf{q}_n - \mathbf{q}_{n-1}\|^2 \tag{3.9}$$

$$E_{bend}(\mathbf{q}_n) = \|\mathbf{q}_{n-1} - 2\mathbf{q}_n + \mathbf{q}_{n+1}\|^2 \tag{3.10}$$

An alternative to the backward difference is a forward difference:

$$E_{elast}(\mathbf{q}_n) = \|\mathbf{q}_n - \mathbf{q}_{n+1}\|^2 \tag{3.11}$$

or a centered difference is also possible:

$$E_{elast}(\mathbf{q}_n) = \frac{\|\mathbf{q}_{n-1} - \mathbf{q}_{n+1}\|^2}{2} \tag{3.12}$$

In [109] it was pointed out that Eq. (3.9) is made under the assumption that the vertices of the active contour model are evenly spaced. As this may not always be the case, it was proposed to subtract the elasticity term from the vertices' average distance $\|\bar{\mathbf{d}}\|$. Unlike Eq. (3.9) the proposed energy expression Eq. (3.13) is not larger for more distant vertices, resulting in a more evenly spaced vertex distribution avoiding a possible contraction of the snake.

$$E_{elast}(\mathbf{q}_n) = \|\bar{\mathbf{d}}\| - \|\mathbf{q}_n - \mathbf{q}_{n-1}\|^2 \tag{3.13}$$

When the above elasticity and bending measures are calculated relative to the initial hypothesis contour $\mathbf{v}_0$ (corresponding to the discrete polygon $\mathbf{P}$) a similarity constraint is introduced [110]:

$$\varepsilon_{elast}(\mathbf{v}(s)) = (\mathbf{v}'(s) - \mathbf{v}'_0(s))^2 \tag{3.14}$$

$$\mathcal{E}_{bend}(\mathbf{v}(s)) = (\mathbf{v}''(s) - \mathbf{v}_0''(s))^2 \tag{3.15}$$

The similarity constraint is flexible in that it implicitly models the properties of the hypothesis. If the hypothesis was constant in its local properties then the similarity favours contours with the same constancy. This is why the similarity constraint results in a smooth contour without explicitly modelling smoothness: through a smooth initial hypothesis.

To explicitly model the smoothness of a contour WILLIAMS suggested a restriction of the local curvature [109]:

$$E_{smoothness}(\mathbf{q}_n) = \cos^{-1} \frac{(\mathbf{q}_n - \mathbf{q}_{n-1})(\mathbf{q}_{n+1} - \mathbf{q}_n)}{\|\mathbf{q}_n - \mathbf{q}_{n-1}\|\|\mathbf{q}_{n+1} - \mathbf{q}_n\|} \tag{3.16}$$

A smoothness constraint that is based on the direction of the image gradient rather than on the contour's geometry is presented in the following section (Eq. 3.25).

3.1.2.2 Image Based Characteristics (External Image Energy Terms)

The external image energy of the classic active contour model [74] involves forces due to image intensities, edges and terminations of line segments, the first of which being the simplest useful image functional:

$$E_{intens}(\mathbf{q}_n) = I(\mathbf{q}_n) \tag{3.17}$$

A simple, discrete edge functional is the intensity gradient (cf. section 5.2.2.1):

$$E_{edge}(\mathbf{q}_n) = -\|\nabla I(\mathbf{q}_n)\|^2 \tag{3.18}$$

where the negative sign produces low energy values for high gradient values. Squaring the gradient narrows the edge response.

In order to find terminations of line segments the curvature of level lines in a slightly smoothed image I_σ can be used to formulate an $E_{term}(\mathbf{q}_n)$.

To form the image energy the terms were integrated using a weighted sum:

$$E_{image}(\mathbf{q}_n) = w_{intens}E_{intens}(\mathbf{q}_n) + w_{edge}E_{edge}(\mathbf{q}_n) + w_{term}E_{term}(\mathbf{q}_n) \qquad (3.19)$$

Depending on the sign of w_{intens} the snake is attracted to either dark or bright pixel intensities. The absolute values of all three coefficients control the relative influences of the energy terms, modelling the predominant image characteristics of the object.

As an alternative to the intensity functional of eq. (3.17), a simple "event detector" was suggested in [111] using a threshold operator with a user defined threshold value θ:

$$E_{thresh}(\mathbf{q}_n) = \begin{cases} 0 & : \quad I(\mathbf{q}_n) < \theta \\ I(\mathbf{q}_n) - \theta & : \quad I(\mathbf{q}_n) \geq \theta \end{cases} \qquad (3.20)$$

The extensive body of research upon image edges has led to the application of a number of different edge detectors in active contours. In addition to the simple gradient operator in Eq. 3.18, gradients with a specified direction are applied [99, 110], with one specific form being the gradient perpendicular to the contour [81]. Among the applied standard edge detectors are the Sobel [72, 100, 105], Canny [105, 108] and Deriche [103] operators. A combination of different 2D Gabor filter responses is proposed in [107].

Of particular relevance to the processing of image sequences is the exploitation of image motion. It seems however, that the results of the extensive research in related subjects such as optical flow have not been integrated into active contours. A motion detector based on the gradient of a difference image D_t was presented in [100]:

$$E_{motion}(\mathbf{q}_n) = |\nabla D_t(\mathbf{q}_n)| \qquad (3.21)$$

where

$$D_t(\mathbf{q}_n) = |I_t(\mathbf{q}_n) - I_0(\mathbf{q}_n) - \bar{I}| \tag{3.22}$$

Here I_t is the current image, I_0 the start image and $\bar{I}$ the mean intensity difference within a reference region to compensate for global intensity variations.

In [112, 113], an inflation or *balloon* force for (closed) active contours was suggested:

$$E_{balloon}(\mathbf{q}_n) = w_{balloon}\mathbf{n}_n \tag{3.23}$$

Where $\mathbf{n}_n$ is a normal unitary vector at $\mathbf{q}_n$ enforcing an expansion of the contour point in direction of its normal. Depending on the sign of the weighting factor $w_{balloon}$ the active contours expands or contracts.

This expansion force, like the classic active contour model in general, is based only on the boundary characteristics of a shape, disregarding the enclosed region pixels even when the contour is closed. Several approaches have been developed to constrain the model's shape to its enclosed region homogeneity, some involving a philosophy similar to that on which template matching is based upon (cf. sections 5.2.1.1 and 5.2.2.1). The most important approach of *statistical snakes* or *active region models* was developed by IVINS and PORILL [114]. Active region models start from a user-defined homogenous template region (or seed region) whose mean μ and variance σ are computed. The contour then grows with the help of an inflation or *pressure* force until it encounters pixels whose intensities change the variance of the regions intensity significantly. The pressure force is defined by

$$E_{pressure}(\mathbf{q}_n) = \mathbf{n}_n \left(\frac{\|I(\mathbf{q}_n) - \mu\| - k\sigma}{k\sigma} \right)^2 \tag{3.24}$$

where k is the constant defining the significance of a change in variance. The pressure force is normalised by the scaling term $(k\sigma)^2$. This scheme is equivalent to weighting the balloon force by the mean pixel intensity at each boundary pixel.[115]

A similar approach has been developed in [116] where a normalised correlation

criterion was proposed, measuring the differences of grey-levels in the current region and the template region.

The above section introduced a general smoothness constraint based on local angles of the contour. GEIGER conversely modelled smoothness based on the direction of the image gradient:

$$E_{smoothness}(\mathbf{q}_n) = |\phi_n - \phi_{n-1}| \tag{3.25}$$

$$\phi_i = \tan^{-1} \frac{\nabla_y \mathbf{q}_i}{\nabla_x \mathbf{q}_i} \tag{3.26}$$

3.1.2.3   Specific Characteristics (External Constraint Energy Terms)

Specific object characteristics in the classic snake model can only be introduced by allowing the user to attach springs between vertices of the contour and fixed positions in the image plane:

$$E_{spring}(\mathbf{q}_n) = -w_{spring}(\mathbf{q}_n - \mathbf{x})^2 \tag{3.27}$$

With a positive value of the spring constant w_{spring} this term attracts the vertex $\mathbf{q}_n$ to a point $\mathbf{x} = (x, y)$ in the image. A negative sign repels the vertex from x in which case the constraint is called volcano force.

However, the need for more specific knowledge particularly on shape has been recognised [117] and led to several inclusions of active contour models.

In [100, 118] the local angle φ_n in the 3-vertex polygon of $\mathbf{q}_{n-1}, \mathbf{q}_n, \mathbf{q}_{n+1}$ was used as a measure of local shape. Together with specific shape constraints such as $\varphi_n = \frac{N-1}{N}\pi$ to model a circle the possible shapes of the active contour were restricted to specific shapes.

The local angle was also used in [118, 119] to attract two closed snakes towards an object boundary. One active contour was initialised on the outside of the object and one on the inside. This set-up improved the evolution of the active contour and helped in finding the optimal solution.

Several other papers address the ability of the active contour to change its shape during the usually iterative optimisation process. BERGER *et al.* in [102] addressed the snake's size which can grow by inserting vertices dynamically. The so-called geodesic active contour presented in [120–122] is even capable of changing its topology, allowing for a rough initialisation.

These approaches reduce the active contour's selectivity with respect to the shape to make it more flexible during the optimisation. Consequently the desired shape of an object cannot be specified precisely which is a drawback where challenging image material is involved.

An approach to detail the desired shape was presented in [97]. Human heads were modelled by using different sets of snake weights on different sections of the contour, corresponding to chin, ears, and hair.

The most complex approach to include not only geometric constraints but to generally represent explicit *a priori* knowledge about specific objects is the grammatical active contour. The novel approach presented in this thesis is partially based on this approach, the details of which are presented in the following.

In [123, 124] OLSTAD proposes a grammatical framework for encoding structural information on the object to be detected. The grammatical object description is integrated into the energy minimising procedures of an active contour model.

Normally syntactic approaches are applied to decompose or classify contours (cf. section 3.2.2.2) *after* they have been segmented. The image information is not utilised in the structural processing of the contour. Incorporating syntactic processing into an active contour affords the possibility of combining both image features and detailed structural constraints in contour segmentation.

One fundamental idea of the grammatical active contour is to specify local properties of an object's contour, that is to use a different energy function for each vertex instead of just one global function. The grammatical description of a contour can then be performed through a string (concatenation) of characters (terminals), such as

aaaaaabbbbbbcbbb

Table 3.1: Example grammatical contour description

where each character stands for one vertex. The letters represent the different, usually external, energy functions E^a_{ext}, E^b_{ext}, and E^c_{ext}.

A hard specification of a characteristic's location such as, "vertex 5 has an angle of 60 degrees", however contradicts the active contour's ability to account for variabilities. A formal grammar offers a representation where the exact location does not have to be specified, but can be left as an additional dimension in the optimisation process.

Formally a (context-free) grammar is defined as a quad-tuple

$$(N, \Sigma, P, S)$$

where

- N is a finite set of non-terminals,

- Σ a finite set of terminals,

- P a finite set of production rules, and

- S the starting symbol of the grammar.

A grammar $G = (N, \Sigma, P, S)$ generates a formal language $L(G)$ which is a set of all strings that can be produced by G. In restricting $L(G)$ to languages that can be described by regular expressions R, OLSTAD introduces a flexibility to the contour description. The basic production rule of concatenation of two regular expressions R_1 and R_2

$$R_1 R_2 = \{st | s \text{ is accepted by } R_1 \text{ and } t \text{ is accepted by } R_2\}$$

ultimately allows for a concatenation of terminals resulting in rigid expressions (cf. Table 3.1). Additional rules however, allow for variation through the union, defined as

$$R_1 \cup R_2 = \{s|s \text{ is accepted by } R_1 \text{ or } R_2\}$$

as well as the KLEENE closure

$$R* = \bigcup_{i=1}^{\infty} R_i$$

allowing for an arbitrary repetition of patterns. Applying the closure operator to the example in Table 3.1 yields

$$\alpha = \text{a*b*c*b*}$$

This regular expression does not specify the exact location of each vertex property anymore. Contours which are segmented based on this description can comply with any of the possible strings as for example

abbbbbbbbbbbcbbb

aaaaaabbbbbcbbb

aaabbbccccbbbbb

Unlike active contours with only one global energy function (corresponding to the expression a*), each vertex can obtain up to $|\Sigma|$ different states. To find the optimal contour in the resulting state space is the task of the optimisation procedure. Olstad integrates a finite-state machine based pattern-matching algorithm into a dynamic programming optimisation (cf. section 3.1.3.2). Effectively this algorithm parses the regular expression, generates a state space containing all possible strings and calculates the overall energy for each of these alternatives.

While the terminals of the grammar could in theory represent any internal or external energy function, Olstad proposes a set of constraint functions which are combined to form a particular external energy function. Image-based constraints are

$$E_{Value}(\mathbf{q}_n) = \nu_s(I(\mathbf{q}_n) - V_0) \tag{3.28}$$

$$E_{InterAvg}(\mathbf{q}_n) = \nu_s(\bar{I}(\mathbf{q}_n, \mathbf{q}_{n+1}) - V_0) \tag{3.29}$$

where $\bar{I}(\mathbf{q}_n, \mathbf{q}_{n+1})$ is the average pixel intensity of all pixels connecting $\mathbf{q}_n$ and $\mathbf{q}_{n+1}$ on a straight line. V_0 is a user defined threshold parameter.

Constraints based on the geometry of a contour are

$$E_{ImageAngle}(\mathbf{q}_n) = f(\angle(\mathbf{q}_n, \mathbf{q}_{n+1})) \tag{3.30}$$

$$E_{Angle}(\mathbf{q}_n) = \nu_e(\varphi_n) = \nu_e(\angle(\mathbf{q}_{n-1}, \mathbf{q}_n, \mathbf{q}_{n+1})) \tag{3.31}$$

$$E_{Movement}(\mathbf{q}_n) = f(\|\mathbf{q}_n - \mathbf{p}_n\|) \tag{3.32}$$

Some of the above functions include weighting functions $\nu(x)$ affording an uncertainty by smoothing the function arguments:

$$\nu_s(x) = 1 - e^{-\sigma/x}, \qquad \sigma = -\ln\frac{1}{2}\sqrt{x} \tag{3.33}$$

$$\nu_e(x) = e^{\beta x} \tag{3.34}$$

where β controls the growth of the exponential function.

3.1.3   Optimisation

Active contour models are considered to be active because the usually iterative minimisation of its energy function causes the model to change dynamically. The slithering movement of the contour during the iterations has lead to the nickname *snakes*. The deformation of the active contour from the initial estimate to its final shape is performed by an energy minimising optimisation process. In the following the classic optimisation technique as well as other techniques applied for active contour models are reviewed.

3.1.3.1   Variational Approach

The classic model is embedded in an Euler-Lagrangian setting, using variational calculus in order to derive a differential equation solved by an iterative minimisation technique using sparse matrix methods.

However, this variational approach does not guarantee global optimality of the solution and requires estimates of higher order derivatives of the discrete data. Moreover, hard constraints cannot be directly enforced unless the constraints are differentiable, in which case higher-dimensional spaces are required for more unknowns. Given a desired constraint term such as a minimum inter-vertex distance, it can only be enforced by increasing the associated weighting, which will force more effect on this constraint, but on the cost of other terms. Further disadvantages of the variational approach are the numerical instability and the tendency for vertices to bunch up on strong image features. [115]

3.1.3.2   Dynamic Programming

To overcome the problems related to the variational optimisation approach presented in the previous section, in [108] AMINI et al. have proposed dynamic programming as an approach to minimise the energy of active contour models. Their approach allows the introduction of hard constraints directly and in a straightforward manner while at the same time ensuring a globally optimal solution with respect to the search space. Numerical stability is ensured by moving the contour vertices on a discrete grid without any approximation requirements.

In general dynamic programming [125] is an optimisation tool with applications beyond the segmentation of contour points for which it has been applied in various contexts [16, 20, 126–128]. The optimisation problem is viewed as a multi-stage decision process, the basic steps of which can be identified as

1. decompose an optimisation problem into smaller sub-problems,

2. find and store the optimal solution for each of the sub-problems,

3. select the solution for a sub-problem, if the sub-problem becomes part of the overall solution

Through this strategy the dynamic programming bypasses local minima as it is embedding the minimisation problem in a family of related problems. With the active contour model this is achieved by replacing the minimisation of the total energy measure by the problem of minimising a number of sub-energies in the form:

$$E_{snake}(\mathbf{Q}) = \min_{dp} E_{snake}(\mathbf{C_0}, \mathbf{C_1}, \ldots, \mathbf{C_{N-1}}) \tag{3.35}$$

$$= E_0(\mathbf{C_0}, \mathbf{C_1}, \mathbf{C_2}) + E_1(\mathbf{C_1}, \mathbf{C_2}, \mathbf{C_3}) + \cdots$$

$$+ E_{N-3}(\mathbf{C_{N-3}}, \mathbf{C_{N-2}}, \mathbf{C_{N-1}}) \tag{3.36}$$

$$= \sum_{i=0}^{i=N-3} E_i(\mathbf{C_i}, \mathbf{C_{i+1}}, \mathbf{C_{i+2}}) \tag{3.37}$$

where each variable is allowed to take only m possible values and

$$E_i(\mathbf{C_i}) = w_{int} E_{int}(\mathbf{C_i}) + w_{ext}(w_{image} E_{image}(\mathbf{C_i}) + w_{con} E_{con}(\mathbf{C_i})) \tag{3.38}$$

In other words, the problem of segmenting the complete contour is decomposed into finding a number of smaller optimal contour segments. As it is not known *a priori* which of the possible contour segments will be part of the globally optimal solution, for each contour candidate its optimal contour segment has to be determined and stored until

in a final backtracking step the optimal segments are connected to form the optimal contour.

The complete dynamic programming algorithm shown in Table 3.2 is based on [108] and [72]. It is used throughout this thesis wherever traditional active contours are applied and forms a basis for the presented novel approach of fuzzy active contours. The central element of the algorithm is the recurrance relation

$$S_i(\mathbf{C_i}) = w_{int}E_{int}(\mathbf{C_i}) + w_{ext}(w_{image}E_{image}(\mathbf{C_i}) + w_{con}E_{con}(\mathbf{C_i})) + \min_{\mathbf{C_{i-1}}} S_{i-1}(\mathbf{C_{i-1}}) \quad (3.39)$$

realising the determination of optimal contour segments. Its realisation $S(n, m)$ in Table 3.2 represents the minimal energy level that is possible for the vertices $0, \ldots, n$ if the nth vertex is the candidate $c_{n,m}$. $T(n, m)$ in the algorithm holds the index k ($k = 0 \ldots M - 1$) that minimises the expression and thus points to the optimal predecessor of the candidate $c_{n,m}$. After all vertices have been processed, the new boundary is obtained by tracing back the pointers, beginning with the candidate that has a minimal $S(N - 1, m)$ value. In most applications the algorithm is repeated until the change of the total active contour's energy ΔE_{snake} is smaller than a prescribed threshold θ.

Typically E_{int} is calculated over the vertices i and $i - 1$ while E_{con} requires information from the vertices i, $i - 1$ and $i - 2$. Consequently in an open contour for the first vertex only E_{image} can be calculated and pointers to preceding candidates do not exist (cf. Table 3.2, lines 1 and 2). The optimisation of the second vertex involves E_{image} as well as E_{int} (lines 3 to 5). The remaining vertices of an open contour as well as all vertices of a closed, periodic contour operate on all energy terms (lines 6 to 9).

As an example Figure 3.2 visualises the segmentation of a contour using dynamic programming. Circles denote the candidates of the active contour. The arrows indicate the pointers calculated during the first phase of the dynamic programming algorithm. Starting at the upper right candidate, which is assumed to have the minimal energy value, backtracking follows the red arrows, connecting optimal candidates (grey circles) for each vertex. Optionally interpolating these vertex positions results in a smooth contour (blue line).

1. **for all** m

2. $S(0, m) = w_{ext} \, w_{image} E_{image}(\mathbf{c}_{0,m})$

3. **for all** m

4. $S(1, m) = \min_k \left[w_{int} E_{int}(\mathbf{c}_{1,m}) + w_{ext} \, w_{image} E_{image}(\mathbf{c}_{1,m}) + S(0, k) \right]$

5. $T(1, m) = k^{\min}$

6. **for** $n = 2 \ldots N - 1$

7.   **for** $m = 0 \ldots M - 1$

8. $S(n, m) = \min_k \left[w_{int} E_{int}(\mathbf{c}_{n,m}) + w_{ext}(w_{image} E_{image}(\mathbf{c}_{n,m}) \right.$

$$\left. + w_{con} E_{con}(\mathbf{c}_{n,m})) + S(n - 1, k) \right]$$

9. $T(n, m) = k^{\min}$

10. $E_{snake} = \min_m S(N - 1, m)$

11. $\mathbf{q}_{N-1} = \mathbf{c}_{N-1, m^{min}}$

12. $m^{min} = T(N - 1, m^{min})$

13. **for** $n = N - 1 \ldots 1$

14. $\mathbf{q}_n = \mathbf{c}_{n, m^{min}}$

15. $m^{min} = T(n, m^{min})$

16. **repeat from** 1. **until** $\Delta E_{snake} < \theta$

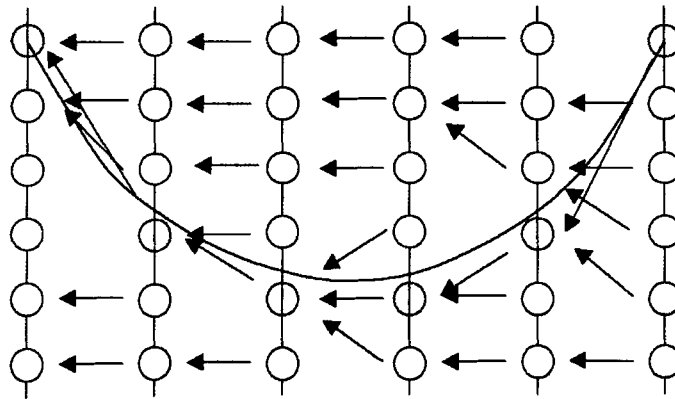Table 3.2: Dynamic programming algorithm for a non-cyclic active contour.

Figure 3.2: Backtracing optimal pointers (red) result in a contour made up of optimal vertices (grey). These may subsequently be used to interpolate a smooth contour (blue).

3.1.3.3 Other Optimisation Strategies

Another optimisation strategy allowing for hard constraints is the greedy algorithm [109], which is regarded as being stable and efficient. The algorithm however does not guarantee to find the global optimum. Also the selection of appropriate thresholds and weights introduces problems in the application of the approach.

Simulated annealing is a stochastic relaxation technique which is based on the physical process of annealing a metal. At high temperatures the atoms are randomly distributed. With decreasing temperatures they arrange in a crystalline state minimising their energy. In [104] this technique was applied to active contours, though the computational demands proved to be very high.

3.2 Survey of Fuzzy Logic in Image Processing

Where uncertainties originate from sources such as the vagueness of linguistic expert knowledge (cf. section 2.2) fuzzy logic is an appropriate approach to both represent and process such uncertainties.

3.2.1 Relevant Basics of Fuzzy Logic

The basics of fuzzy logic is the fuzzy set theory founded by ZADEH [129, 130]. Fuzzy set theory '*in the last two decades has developed along two lines:*

1. *As a formal theory which, when maturing, became more sophisticated and specified and was enlarged by original ideas and concepts as well as by 'embracing' classical mathematical areas such as algebra, graph theory, topology, and so on by generalizing (fuzzifying) them.*

2. *As a very powerful modeling language, that can cope with a large fraction of uncertainties of real-life situations. Because of its generality it can be well adapted to different circumstances and contexts. (...)'*[131]

The basic element of fuzzy set theory is the extension of classic or crisp sets to *fuzzy sets* while at the same time providing appropriate operations on these continuous sets.

This concept provides the basis for an extension of the traditional Boolean logic to a continuous or *fuzzy logic*. Here logic variables are defined as fuzzy sets while the operations defined for fuzzy sets allow for a definition of fuzzy logic operations such as logical AND and OR. Now it is possible to formulate *fuzzy logic rules* and to draw conclusions.

While the continuous representation of input variables is closer to the real world than the binary logic it is still not close enough to the human representation of vague information. Here the key point is language, which is integrated into fuzzy logic through *linguistic variables*. This concept allows for the mapping of linguistic terms to fuzzy sets, consequently allowing for a formulation of logic rules in a linguistic form as *linguistic rules*.

As the next more complex application of fuzzy rules *fuzzy inference systems* can process an input through a collection of fuzzy rules, the rule or knowledge base. Unlike such systems in Boolean logic an initial fuzzification step is needed to relate the usually crisp input values to fuzzy sets which can be handled by the system. By definition the result of the subsequent inference of the fuzzy rules is a fuzzy conclusion. Should a non-fuzzy output be required, a defuzzification is performed to translate the fuzzy result back to a crisp value.

In the following the above mentioned elements of fuzzy set theory and fuzzy logic are briefly summarised to define a notation and to make this thesis reasonably self-contained. For details consider the many available text books such as [131–140].

3.2.1.1   Fuzzy Sets

Sets in general can be regarded as mathematical abstractions of objects in the real world. Set theory provides such abstractions as well as operations to process these sets.

A classical or crisp set $\mathcal{A}_c$ can be defined as a collection of objects, or elements, $x \in \mathcal{X}$ where $\mathcal{A}_c \subseteq \mathcal{X}$. The definition of $\mathcal{A}_c$ can be performed by either listing its elements, for example

$$\mathcal{A}_c = \{1, 2, 3, 4, 5\}, \tag{3.40}$$

by stating conditions for membership

$$\mathcal{A}_c = \{x | x \le 5, x \in \mathbb{N}\} \tag{3.41}$$

or by using the characteristic function $f_{\mathcal{A}_c}$, in which 1 indicates membership and 0 non-membership respectively:

$$f_{\mathcal{A}_c} = \begin{cases} 1 & : & x \le 5, x \in \mathbb{N} \\ 0 & : & x > 5, x \in \mathbb{N} \end{cases} \tag{3.42}$$

This dual membership is characteristic for a crisp set. A fuzzy set $\mathcal{A}$ however, is characterised by a membership function μ_A, which assigns to each element $x \in \mathcal{X}$ a degree of membership where $\mu_A(x) \in [0, 1], \mu_A(x) \in \mathbb{R}$:

$$\mathcal{A} = \{(x, \mu_A(x)) | x \in \mathcal{X}\} \tag{3.43}$$

An application for a fuzzy set is the definition of a fuzzy number. Figure 3.3 illustrates such a (discrete) fuzzy number l, with mean value l_0 and spread s, defined by $\mu(l) > 0 \; \forall \; l \in \{l_0 - s, \ldots, l_0 + s\}$. An example "approximately 3" can be defined by

$$\mathcal{A} = \{(1, 0.3), (2, 0.6), (3, 1.0), (4, 0.6), (5, 0.3)\} \tag{3.44}$$

Note that for convenience those elements with $\mu_A(x) = 0$ are normally not listed.
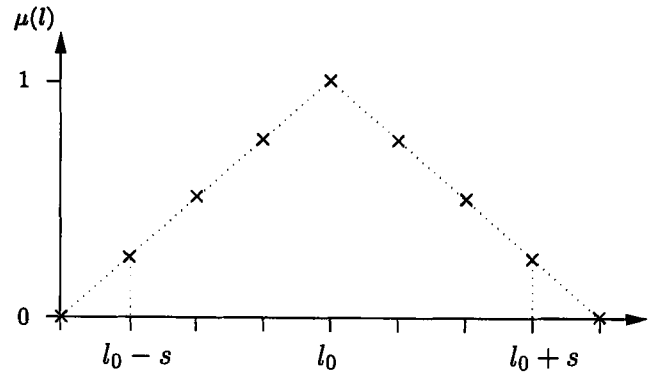
Figure 3.3: A triangular fuzzy number.

3.2.1.2 Fuzzy Rules

Logics can be characterised by their truth values, operators and reasoning procedures. In Boolean logic, truth values can be either 0 (false) or 1 (true). Based on these truth values operators are defined via truth tables. Two familiar examples for two statements A and B being

A	B	$\wedge$	$\vee$
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	1

defining the operators AND ($\wedge$) and OR ($\vee$). There are situations however, where two truth values are not sufficient and where truth tables cannot easily be assigned a name to reflect its function [131]. As a solution fuzzy logic regards both the truth values and the statements as fuzzy sets (or more precisely as linguistic variables which are presented below).

Let $v(A)$ be a fuzzy set representing the truth value of the statement A with

$$v(A) = \{(x, \mu_A(x)) | x \in [0,1]\}, \qquad (3.45)$$

then the truth value TRUE can be defined as a fuzzy set

$$v(\text{TRUE}) = \{(0.5, 0.6), (0.6, 0.7), (0.7, 0.8), (0.8, 0.9), (0.9, 1.0), (1.0, 1.0)\} \quad (3.46)$$

With the negation operation defined as

$$\neg v(A) = v(\text{NOT } A) = 1 - v(A) = \{(1 - x, \mu_A(x) | x \in [0, 1]\}, \quad (3.47)$$

the truth value FALSE becomes

$$v(\text{FALSE}) = \{(0.0, 1.0), (0.1, 1.0), (0.2, 0.9), (0.3, 0.8), (0.4, 0.7), (0.5, 0.6)\} \quad (3.48)$$

The logical operations AND and OR are normally defined as

$$v(A) \wedge v(B) = v(A \text{ AND } B) = \{(x, \min[\mu_A(x), \mu_B(x)])\} \quad (3.49)$$

and

$$v(A) \vee v(B) = v(A \text{ OR } B) = \{(x, \max[\mu_A(x), \mu_B(x)])\} \quad (3.50)$$

respectively. Similarly the operator $\Rightarrow$ (implies) is defined as

$$v(A) \Rightarrow v(B) = v(A \Rightarrow B) = \neg v(A) \vee v(B) = \{(x, \max[1 - \mu_A(x), \mu_B(x)])\} \quad (3.51)$$

providing the basis for reasoning procedures such as the common *modus ponens*

$$(A \wedge (A \Rightarrow B)) \Rightarrow B$$

It is applied in reasoning to come from a premise (A is TRUE) and an implication (IF A THEN B) to a conclusion (B is TRUE). Using fuzzy logic allows *approximate reasoning* where

1. the statements as well as

2. the conclusion may be fuzzy.

The elements of fuzzy logic described so far provide the formal basis for an intuitive interface to human reasoning by introducing fuzzy inputs (statements) and output (conclusions).

To make the transition from formal rules such as

$$\text{IF } A \text{ is } 0.6 \text{ AND } B \text{ is } 0.8 \text{ THEN } C \text{ is } 0.9$$

to a more intuitive representation of knowledge in the form of linguistic rules such as

$$\text{If the pixel is bright then noise is very likely}$$

it is necessary to integrate elements from the human language into fuzzy logic. The concept of linguistic variables provides this integration and is summarised in the following.

3.2.1.3 Linguistic Variables

The first step in the integration of linguistic terms into fuzzy logic is an analysis of the possible interpretations the concept of the fuzzy membership allows. The membership plays a central role in fuzzy systems as it represents the connecting element between the real world and the numerical world of mathematics and logic. Each application may require a different view on this relation. TIZHOOSH gives an overview about the different interpretations of the membership in the literature:

- **Membership as similarity** When classifying objects through the definition of a prototype, memberships reflect the degree of similarity between an object an the prototype. The feature space is used to determine this degree of similarity. Assigning a membership of 0.9 to an object could for example represent the numerical mapping of the linguistic terms "very similar" or "nearly identical".

- **Membership as probability** Membership can also be interpreted as an object's probability of belonging to a certain class. The terms "very likely", "likely" and so on can be represented as fuzzy sets. When stochastic processes cannot be handled by probability theory, these terms may be determined subjectively. In this case using fuzzy sets may be useful to describe such processes. However, fuzzy membership itself must not be confused with probability (on the differences consider for example [131]).

- **Membership as approximation** In some applications the membership degree reflects the quality of an approximation of a measured value and the real or reference value.

- **Membership as intensity** A different interpretation of membership is the intensity with which the property of a set $\mathcal{X}$ is true for an element $x \in \mathcal{X}$. A good example is the brightness for which terms such as "very bright", "dark" and so on can be found.

These interpretations illustrate the usefulness of a linguistic representation of membership degrees. ZADEH formulated the motivation as follows:

> '*In retreating from precision in the face of overpowering complexity, it is natural to explore the use of what might be called* linguistic variables, *that is, variables whose values are not numbers but words or sentences in a natural or artificial language. The motivation for the use of words or sentences rather than numbers is that linguistic characterizations are, in general, less specific than numerical ones.*'[130]

The concepts of a linguistic variable formalises the relation between linguistic and numerical values as follows. A linguistic variable is defined by a quintuple

$$(x, \mathcal{T}(x), \mathcal{U}, \mathrm{G}, \mathrm{M})$$

in which

- x is the name of the variable,

- $\mathcal{T}(x)$ is a set of its terms or linguistic values,

- $\mathcal{U}$ the universe of discourse for x, with the base variable u,

- G is a (syntactic) rule generating the term names X, and

- M is a (semantic) rule for associating with each X its "meaning" $M(X)$ which is a fuzzy set on $\mathcal{U}$.

Figure 3.4 illustrates the transition from the linguistic level to the numerical level for the following example. Note that in this example the rules G and M are not formally defined. Both the generation of term names and the definition of associating fuzzy sets are performed subjectively by a human expert. This is the case in many applications and shows where subjectivity is introduced in the otherwise precise definition of a linguistic variable.

Consider a linguistic variable with

$$x = pixel\ brightness$$

$$\mathcal{T}(x) = \{\text{very dark, dark, grey, bright, very bright}\}$$

With $\mathcal{U} = [0, 255]$ the base variable u is the pixel brightness in grey levels. An example for the assignment of a fuzzy set to the term $X = \text{dark}$ is

$$M(\text{dark}) = \{(u, \mu_{\text{dark}}(u))\}$$

where

$$\mu_{\text{dark}}(u) = \begin{cases} \frac{u}{63} & : \quad u \in [0, 63] \\ 1 - \frac{u-64}{126-64} & : \quad u \in [64, 126] \\ 0 & : \quad u > 126 \end{cases}$$
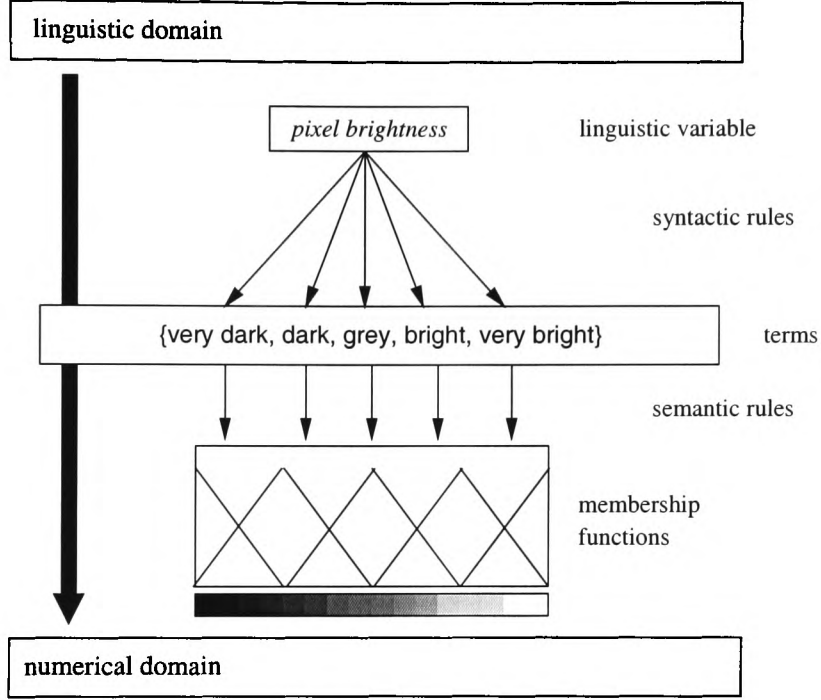

Figure 3.4: Transition from the linguistic domain to the numerical domain through a linguistic variable. (Adapted from [22])

3.2.1.4 Linguistic Rules

Applying the concept of linguistic variables to fuzzy logic rules is straightforward. For each input and output statement a linguistic variable is applied. In the example rule

IF *the pixel* is bright THEN *noise* is very likely

for the input statement *the pixel* the linguistic variable *pixel brightness* is applied, while for the output statement *noise* a linguistic variable *probability* with $\mathcal{T}(x) =$ {not likely, ..., very likely} is defined and applied.

The way such linguistic rules are applied in an inference system is summarised in the following section.

3.2.1.5 Fuzzy Inference Systems

Applying the above elements of fuzzy logic to a practical problem requires an appropriate tool. Fuzzy inference systems (FIS) provide such a tool which is powerful and flexible enough to solve different problems.

Although historically applied mainly to control problems, fuzzy inference systems have proved to be useful in a variety of application areas. The main motivation being the possibility to model systems by exploiting expert knowledge when they cannot be modelled mathematically.

To apply a fuzzy inference system to problems based on numerical rather than fuzzy values, it is necessary to transform or code the numerical input values to the fuzzy domain. They are then processed by inferencing a rule base which holds expert knowledge represented through a collection of (linguistic) fuzzy rules. Finally, the result in the fuzzy domain is transformed back to a crisp numerical value. The coding (fuzzyfication) and decoding (defuzzyfication) are normally necessary as fuzzy inference systems are usually only one component in an otherwise crisp information process.

Figure 3.5 shows the general structure of a fuzzy inference system, the components of which are detailed in the following.
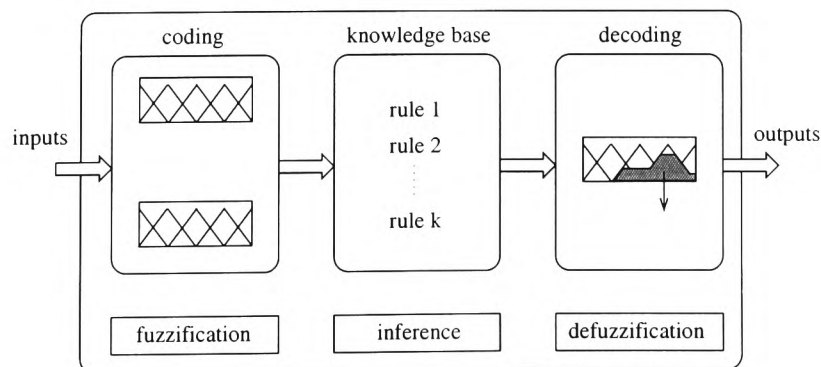

Figure 3.5: General structure of a fuzzy inference system. (Adapted from [22])

Fuzzification — Coding of Inputs

The input of a fuzzy inference system is defined by its numerical (crisp) input variables $u_i \in \mathcal{U}_i, i \in [1, I]$. For each input, a linguistic variable x_i is defined upon creation of the fuzzy inference system. When the system executes the fuzzification procedure, for each input value its memberships to all the linguistic terms of the associated linguistic variable are determined:

$$\forall X_{ji} : \mu_{X_{ji}}(u_i)$$

As an example consider the above defined linguistic variable $x = pixel\ brightness$. If an input grey level is $u_1 = 150$ then the following memberships are calculated as (cf. Figure 3.6): $\mu_{\text{very dark}}(150) = 0$, $\mu_{\text{dark}}(150) = 0$, $\mu_{\text{grey}}(150) = 0.6$, $\mu_{\text{bright}}(150) = 0.4$, $\mu_{\text{very bright}}(150) = 0$. The whole of these memberships represent the fuzzified pixel brightness value 150.
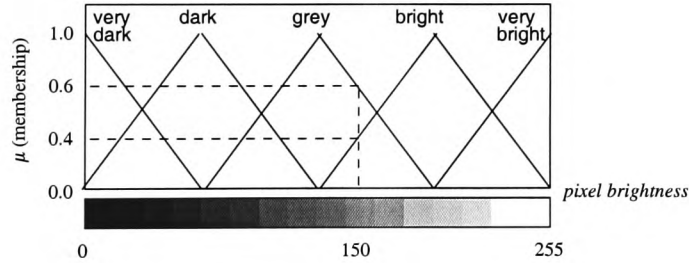

Figure 3.6: Fuzzification example for the linguistic variable *pixel brightness*. (Adapted from [22])

Rule Base — Representing Expert Knowledge

The relation between the inputs of a fuzzy inference system and its outputs is described through a collection of linguistic rules — the rule base, or more general the knowledge base. Upon creation of the FIS, the rules $R_k, k \in [1, K]$ are set up to represent an expert's knowledge of the problem to be solved by the FIS. The formal specification

as well as the complexity of a rule depends on the functionality of the actual implementation of the FIS which is applied. In general simple IF-premise-THEN-conclusion rules are applied in the form

$$R_k : \text{IF } P_k \text{ THEN } C_k$$

A simple example rule base follows:

R_1 : IF *pixel brightness* is dark   THEN *noise* is not likely

R_2 : IF *pixel brightness* is grey   THEN *noise* is likely

R_3 : IF *pixel brightness* is bright THEN *noise* is very likely

Where multiple input variables are involved their dependencies are accounted for through premises composed of sub-premises P_{kl} connected by fuzzy logical operators $\odot_{kl}$:

$$P_k = P_{k1} \odot_{k1} P_{k2} \odot_{k2} \cdots$$

Inference — Reaching a Conclusion

For simple rules such as in the above example it is sufficient to execute an implication operation to come from a premise to a conclusion. A rule base consisting of a number of rules however which may also contain multiple sub-premises, for example

R_1 : IF $P_{11} \lor P_{12}$ THEN C_1

R_2 : IF $P_{21} \land P_{22}$ THEN C_2

requires an inference procedure involving the three stages aggregation, implication and accumulation.

Aggregation. For each sub-premise P_{kl} its truth value is calculated. Operators then aggregate these in order to obtain a single truth value v_k for the whole premise P_k. The aggregation operators correspond to the fuzzy logical operators $\odot_{kl}$ specified in the rule. Normally fuzzy logical AND and OR are applied where the corresponding aggregation operators are min and max respectively. For the above example the following operations are performed:

$$v_1 = v(P_{11} \vee P_{12})$$

$$v_2 = v(P_{21} \wedge P_{22})$$

Implication. The task of this stage is to determine the degree of fulfillment of a conclusion based on the truth values of the premise. For each rule R_k of the same output variable, the truth value of its premise v_k is used to weight the membership function of the linguistic term given in the conclusion C_k. The result is a fuzzy set $\mathcal{Z}_k$. Normally the minimum or the product operator is applied.

For example with

$$v_1 = v(\textit{pixel brightness} \text{ is } \mathbf{dark}) = 0$$

$$v_2 = v(\textit{pixel brightness} \text{ is } \mathbf{grey}) = 0.6$$

$$v_3 = v(\textit{pixel brightness} \text{ is } \mathbf{bright}) = 0.4$$

an implication applying a minimum operator yields

$$\mathcal{Z}_1 = \{(u, \mu_{\mathcal{Z}_1}(u))\} = \{(u, \min(v_1, \mu_{\mathbf{not\ likely}}(u)))\}$$

$$\mathcal{Z}_2 = \{(u, \mu_{\mathcal{Z}_2}(u))\} = \{(u, \min(v_2, \mu_{\mathbf{likely}}(u)))\}$$

$$\mathcal{Z}_3 = \{(u, \mu_{\mathcal{Z}_3}(u))\} = \{(u, \min(v_3, \mu_{\mathbf{very\ likely}}(u)))\}$$

Accumulation. The accumulation stage integrates the fuzzy result of $\mathcal{Z}_k$ of each rule's implication into a single fuzzy result $\mathcal{Z}$ for all rules involving the same output. Normally the maximum operator is applied in the accumulation:

$$\mathcal{Z} = \{(u, \mu_{\mathcal{Z}}(u))\} = \{(u, \max_k(\mu_{\mathcal{Z}_k}(u)))\}$$

Defuzzification — Decoding of outputs

To transform the fuzzy result of the inference back to the numerical domain, that is to a crisp output value, defuzzification is performed.

The design of a fuzzy inferences system includes the definition of output variables $u_o \in \mathcal{U}_o, o \in [1, 0]$. Analogous to the fuzzification, for each u_o a linguistic variable z_o is defined. To obtain a crisp value u_o from the fuzzy result $\mathcal{Z}$ several different methods exist, and selecting the most appropriate depends on the particular application. From a philosophical viewpoint as this stage requires an interpretation of the meaning of all possible fuzzy sets $\mathcal{Z}$ this choice is not always straightforward. Consequently devising new interpretations (that is, defuzzification methods) is subject to continuous research. One practically relevant method is presented here as it is used in this thesis.

The centre-of-area (COA) method, also called centre-of-gravity (COG) method, interprets the centre of the area spanned by the fuzzy set $\mathcal{Z}$ as the best crisp interpretation of this set. The COA is defined as

$$u = \frac{\sum\limits_{u \in \mathcal{U}} u \, \mu_{\mathcal{Z}}(u)}{\sum\limits_{u \in \mathcal{U}} \mu_{\mathcal{Z}}(u)} \tag{3.52}$$

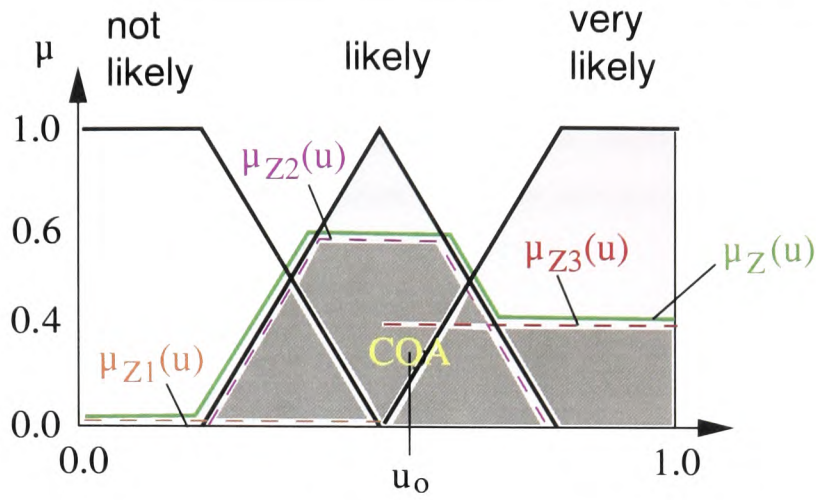To visualise this method consider the example provided in Figure 3.7

Figure 3.7: Defuzzification example applying the centre-of-area method.

3.2.2 Fuzzy Image Processing

Fuzzy image processing as an independent area of research is a relatively young discipline. It seems that there are no comprehensive text books. A first attempt to integrate the many separate theoretical approaches developed over the last two decades is [22]. This book of TIZHOOSH as well as most of the literature surveyed in this project mainly covers low-level methods and local operations such as image filtering. Practically every well-referenced operator developed in (crisp) image processing research has a fuzzy counterpart.

TIZHOOSH defines fuzzy image processing as follows:

'*Fuzzy image processing includes all approaches of digital image processing, where the images, their segments or features of images and their segments are interpreted, described and processed as fuzzy sets. In which way this is performed, depends both on the problem and the particular fuzzy approach.*'

Of particular relevance to this thesis is the question of how fuzzy logic can be utilised to improve the segmentation of vaguely describable objects. Most existing approaches

apply fuzzy logic in low or intermediate levels of image processing with either little or no relation to complete objects. There does not seem to exist a method to describe an object on a high level of abstraction, preferably linguistic, while at the same time using this description in a segmentation procedure. Some relevant approaches are summarised in the following.

3.2.2.1 Object-independent processing

Amongst the few recent medical image processing applications where fuzzy logic is applied are [141, 142] where tumors are to be segmented in MR images. Both articles present a low-level pixel-based segmentation where pixels are classified into tissue classes. The classifiers are trained in a preceding supervised or unsupervised learning stage.

When concerned with segmenting images the subject of (object independent) edge detection has long since attracted the attention of many researchers. An early example for an edge detection method applying fuzzy set theoretic methods is [143]. PAL and KING present the segmentation of wrist bones in X-ray images, where in a first stage fuzzy contrast enhancement and smoothing are applied. To perform these fuzzy operations the grey-level image is transformed to the fuzzy domain by interpreting each pixel as a fuzzy set. The membership values denote the degree of having a brightness level relative to some brightness level. After the fuzzy enhancement operations are performed the image is transformed back to the spatial image domain. The actual edge detection is performed as a local operation in this crisp domain.

Consequently this approach does not fully integrate fuzzy set theoretic operations into an edge detection algorithm. In [144] a fully integrated pixel-based edge detection is presented based on a model of an edge. The model is described explicitly and processed using a fuzzy reasoning approach. An implicit representation of an edge is used in [145], where fuzzy edges are learnt from 6 × 6-pixel templates.

The concept of a fuzzy image utilised in [143] was refined in [132] in that the mem-

bership values of the pixel's corresponding fuzzy sets can be related to any property, depending on the application. This notion was applied in [146] for example. Here the fuzzy property relates to uniform surfaces. Applying fuzzy image processing operations such as smoothing leads to a manipulation of the image with respect to the modelled property, such as smoothing surfaces. While this approach does not merely look at local edges it still defines a general object property and as such it is still independent from a specific object.

3.2.2.2 Object-related processing

The description of object contours utilising fuzzy methods has mainly been studied in the context of analysing *previously segmented* contours. Here shape analysis and contour decomposition are the main subjects to facilitate feature selection and primitive extraction in that context.

PAL *et al.* have presented a measure of the fuzzy "arcness" of a curved line segment x:

$$\mu_{arc}(\mathbf{x}) = \left(1 - \frac{a}{l}\right)^{F_e} \tag{3.53}$$

where a is the Euclidean distance between the two endpoints of x and l is the actual length of the arc segment. As this measure is applied in the fuzzy domain, the formalism contains a fuzzification parameter F_e (cf. [132] or details).

A segmented, binary contour is chain coded and analysed using this arcness measure to distinguish three primitives: a straight line L_l, with l being the line length, a clockwise arc A_μ, with the degree of arcness μ, and a couter-clockwise arc $\bar{A}_\mu$ respectively.

A typical description of a contour resembles for example

$$L_{11}A_{0.86}L_4A_{0.272}L_1\bar{A}_{0.662}L_4A_{0.598}L_7\bar{A}_{0.272}A_{0.765}\bar{A}_{0.816}\bar{A}_{0.272}L_1A_{0.765}$$

This intermediate-level representation is a preceding stage in utilising high-level knowledge to recognise the contour, that is the underlying object. PAL suggests to apply a set of application specific thresholds to the above basic primitives to select a specific set of primitives. These are then used to develop a grammar to recognise the contour, that is its resulting string, syntactically [147].

It is desirable to perform such a syntactic processing with a continuous version of formal grammars. Such fuzzy grammars exist [22, 132] as well as their related mechanisms such as fuzzy automata. A comprehensive bibliography is given in [148]. A recent application example is the recognition of hand-written script [149]. Previously segmented contours are decomposed into sequences of segments of constant curvature. Using a fuzzy syntactical approach these segments are assigned to sub-allographs, which in turn are related to allographs, which is when the recognition is reached. Both sub-allographs and allographs are defined by a fuzzy grammar. Each sub-allograph is a sequence of primitives, each primitive possessing a starting point, an end point as well as one or more characteristic points in between. Features can be derived from these primitives, such as curveness, discontinuity and tilt.

Using fuzzy grammars to incorporate a priori knowledge into the actual segmentation process, similar to OLSTADs approach utilising traditional grammars (cf. section 3.1.2.3), does not seem to exist.

RALESCU et al. [150] defines fuzzy primitives not to analyse previously segmented contours but to actually perform a segmentation step. This approach is based on a group-and-delete technique where an image is binarised into a large number of short line segments. These lines are then linked or erased depending on some criterion. With RALESCU et al. the grouping is based on perceptual organisation. Basic human visual properties are mimicked by the model in recognising straight lines or L-junctions (and the way we distinguish one from the other). The related primitive consists of two straight line segments L_1 and L_2 as well as two constraints on the endpoint distance and the inner angle between these lines (cf. Figure 3.8a). Both constraints are modelled through fuzzy sets reflecting the opinions of an expert on when to consider two lines

near enough to belong to a common line or junction (cf. Figure 3.8b), as well as when the inner angle represents a straight line or a junction (cf. Figure 3.8c).
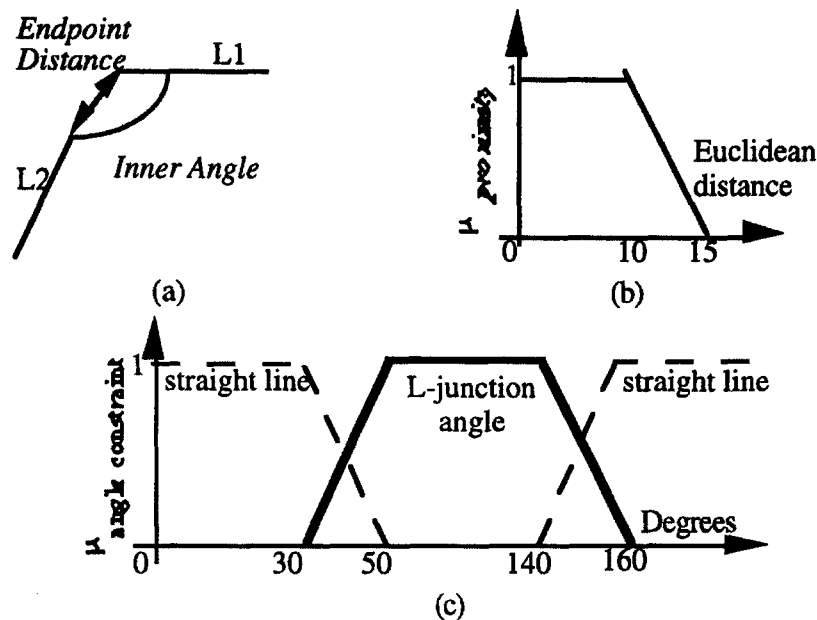

Figure 3.8: Modelling L-junctions: a) example of L-junction (endpoint distance and inner angle); b) proximity membership function; c) L-junction angle membership function (dashed membership functions are for proximity of colinear segments. (from [150])

Chapter 4

A Fuzzy Active Contour

This chapter introduces the concept and presents the theoretical basis of the fuzzy active contour. Experimental results are presented in chapter 5 to demonstrate the validity of the theoretical properties of the fuzzy snake approach and to illustrate the application of the technique to medical imaging sequences.

4.1 Introduction to Fuzzy Active Contours

A new method for representing and tracking of object boundaries is presented in this chapter. The novel concept of fuzzy active contours or fuzzy snakes is developed to allow for the integration of uncertain *a priori* knowledge into an active contour model.

The fuzzy active contour combines the concept of active contours with elements of fuzzy logic where the following benefits of each approach are exploited.

Active contours

- provide a segmentation method which is capable of handling variations in image features and object shape,

- integrate low-level constraints to utilise general object knowledge,

- allow for the integration of high-level constraints to utilise specific object knowledge.

Fuzzy logic offers

- a formal theory which provides tools to handle uncertainties, such as

- linguistic rules which can provide an intuitive, verbal specification of the properties of an object's boundary to utilise high-level expert knowledge;

- appropriate algorithms and interfaces to process linguistic rules and integrate this fuzzy concept into a crisp numerical context.

Analogous to the image energy E_{image} of an active contour, the fuzzy active contour is capable of representing the appearance of an object in the image. Further properties of the object's boundary segments, such as shape, may be represented in a similar way to the function E_{con} in an active contour.

This is achieved by introducing fuzzy energy functions and establishing a linguistic rule base, which describes each of the fuzzy snake's segments. The contour candidate's external energy E_{ext} is then evaluated through fuzzy inference and subsequent defuzzification (cf. section 3.2.1.5).

In [151] deformable models are combined with fuzzy image processing operators to exploit uncertain knowledge on a low level. The proposed fuzzy active contour goes beyond that in providing an interface to high-level uncertain knowledge.

Figure 4.3[1] shows the principal structure of the calculation of the fuzzy energy function with integrated linguistic rules. Similar to traditional active contours (cf. Figure 4.1) salient image-based features are extracted. Geometric features are also calculated providing a basis for the exploitation of *a priori* shape knowledge. Note that in principle any additional constraints may be integrated into the fuzzy active contour.

---

[1]For a brief introduction to the Structured Analysis and Design Technique (SADT) refer to Appendix A.

Image as well as constraint features are fuzzified to allow for their processing in a fuzzy inference machine. This stage evaluates a set of rules which specify the desired properties of the currently processed contour candidate. After defuzzification an external energy value is output. It is compatible with traditional active contours in that low values represent a good compliance of the candidate with the model. Unlike active contours the fuzzy snake's model is partially represented by a set of linguistic rules.

The traditional active contour approach was investigated during this research (cf. chapter 5). One of the findings is that the results can be improved considerably by extracting image features through multiple methods, obtaining multiple image evidences. Consequently the active contour was extended by a weighted sum of edge-, region-, and motion-based operators (cf. Figure 4.2, modules A1–A3).

The fuzzy active contour also exploits multiple image evidences. They are not simply added however, but fully integrated into the overall fuzzy concept (cf. Figure 4.4, modules A1–A3).
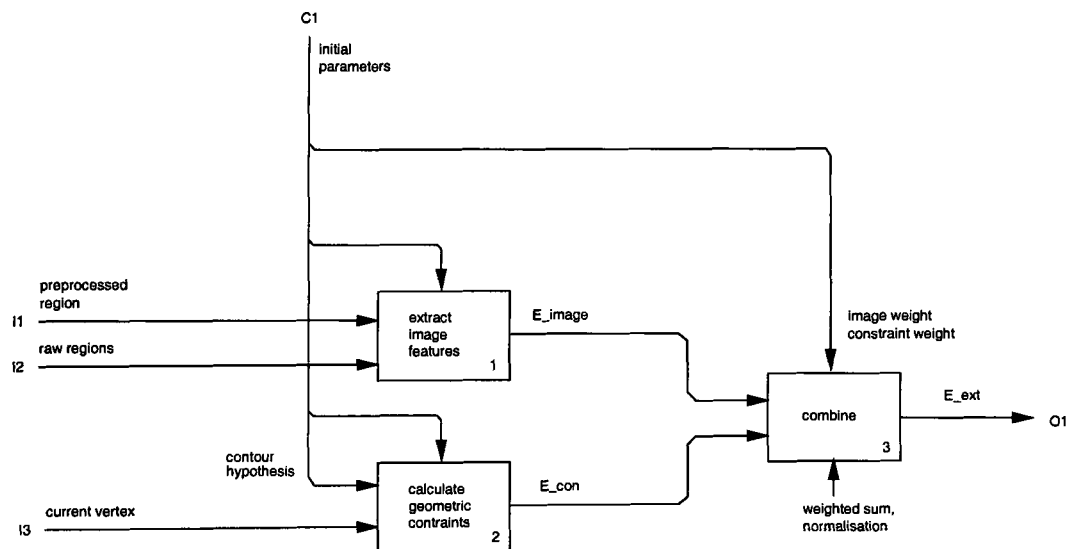
Figure 4.1: Principle structure of external energy calculation for traditional active contours.
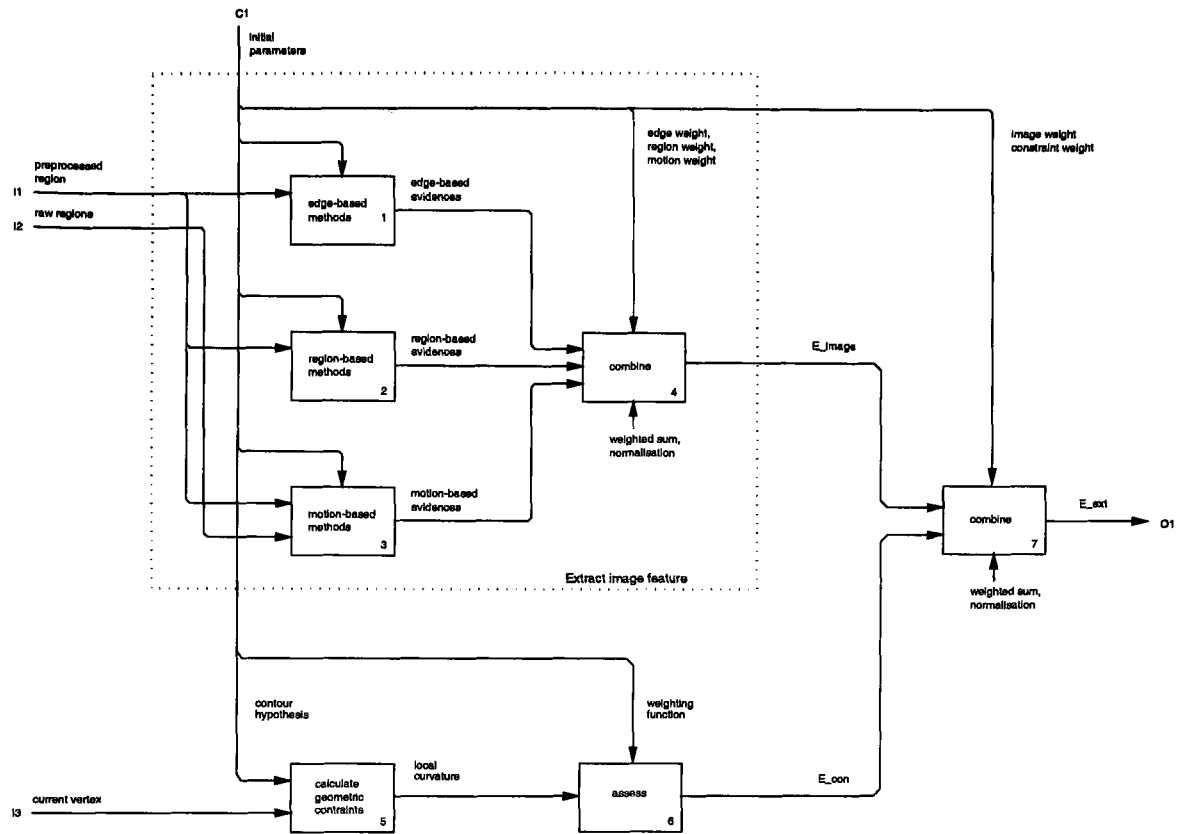
Figure 4.2: Principle structure of external energy calculation for traditional active contours using multiple image evidences.
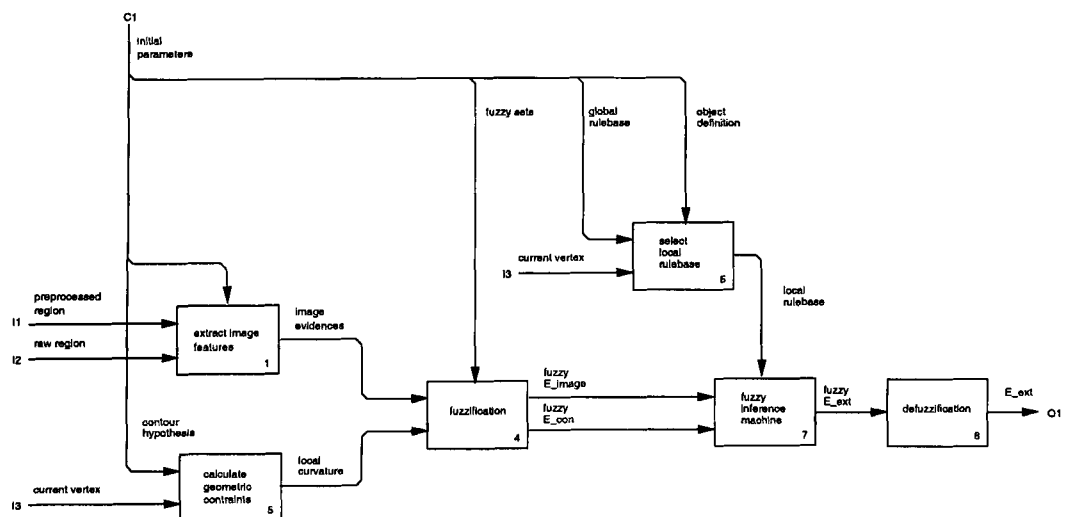

Figure 4.3: Principle structure of external energy calculation for the fuzzy active contour.
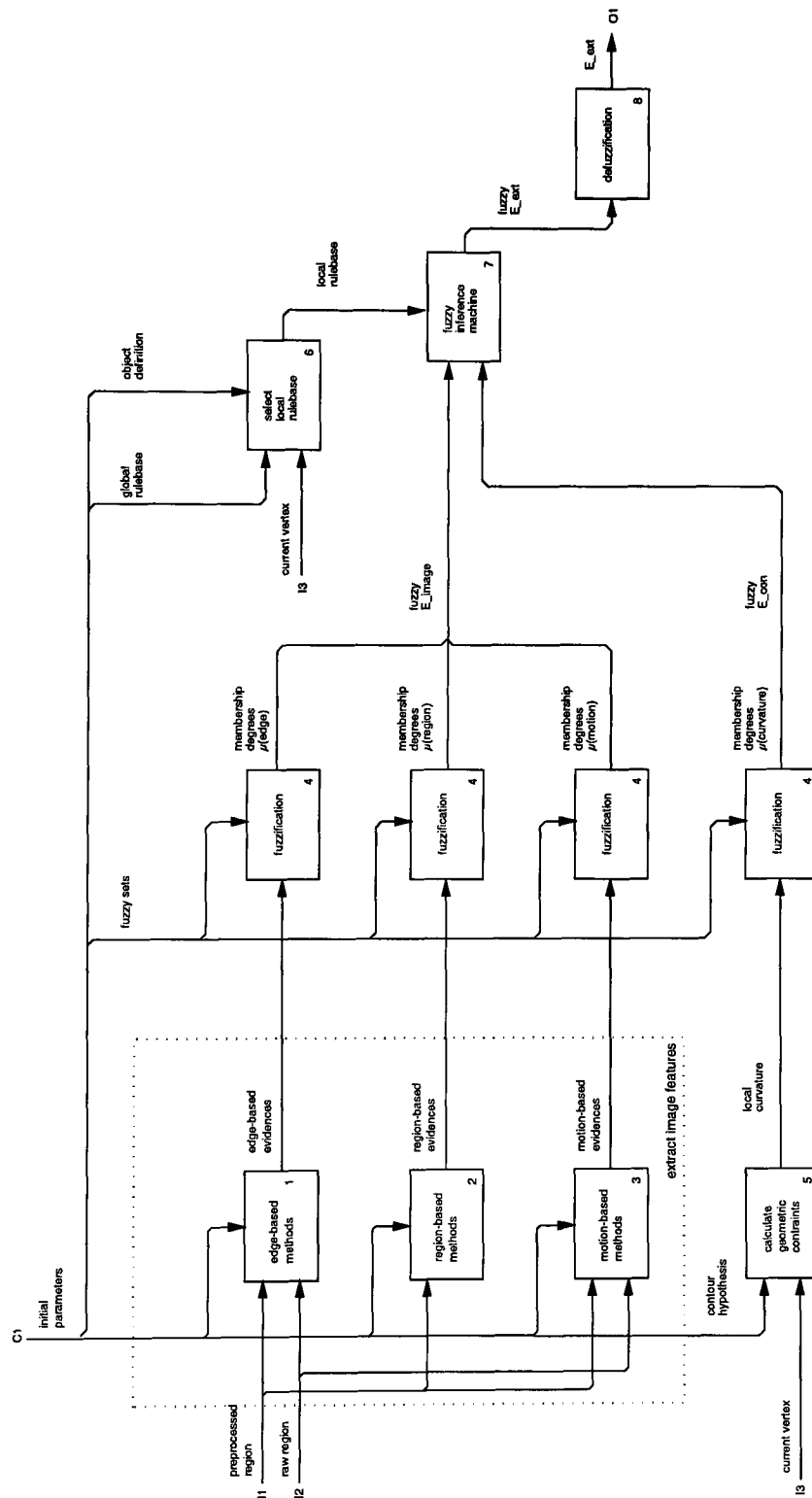
Figure 4.4: Principle structure of external energy calculation for the fuzzy active contour. This diagram also details the use of multiple image evidences.

Also unlike traditional snakes for each segment of an object's boundary a different set of rules may be applied. Furthermore the approximate length of each contour segment may be specified to both improve the segmentation process and to reduce computational complexity.

This approach allows for a formal specification of an object which is very similar to a human verbal description such as

"A medium length dark arc, bending right, followed by a medium length grey line,

a right bending flexible corner and a short rigid line."

The formal description is easily derived from the above, simplifying knowledge engineering as well as the application of the fuzzy active contour. A global set of linguistic rule bases is defined for each application, defining what features make for example a "dark arc" or a "rigid line". As these fuzzy primitives may consist of a variable number of contour vertices, a particular selection method determines which rule base has to be applied for a given contour candidate.

4.2 Fuzzy Contour Model

4.2.1 Multiple Segments

Despite the advances in the development of traditional active contours a fundamental limitation remained in terms of representing *a priori* knowledge concerning more complex objects which were to be detected. Namely that all vertices of the active contour were characterised by the same local energy function, resulting in a single, global description of the object.

In order to overcome this problem, OLSTAD introduced a grammatical description of the snake's energy function (cf. section 3.1.2.3). Figure 4.5 shows an example of a contour, which can be described as a sequence of four different external energy functions E_{ext}^a, E_{ext}^b, E_{ext}^c and E_{ext}^d, represented by the respective terminals a, b, c

and d. The grammatical expression describing such a segmented boundary would be $\alpha = a*b*c*d*$, where $*$ is the closure operator which allows parts of a pattern to be arbitrarily repeated. A pattern-matching algorithm is incorporated within the active contour's energy minimisation process which constrains the resulting contours to only those that comply with the grammatical expression. Three fundamental drawbacks in this algorithm can be identified:

1. The different energy functions do not intrinsically consider inexact *a priori* knowledge.

2. The length of a contour segment cannot be specified, although an approximate value may be known in advance.

3. The algorithm is computationally expensive, since the closure operation generates a large number of possible states in the finite-state-machine based pattern matching.

The fuzzy active contour solves these problems. It retains only the basic principle of such a multi-segment contour model while the philosophy of both its high-level contour description and its algorithmic details are substantially different.
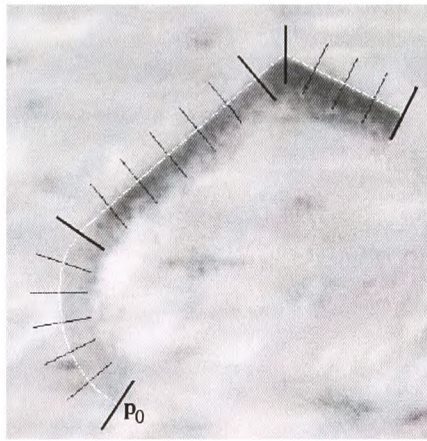

Figure 4.5: Example object. This synthetic boundary demonstrates a multi-segment object contour. The boundary segments have a different shape (local curvature) as well as a different appearance in the image (edge contrast), while all vertices within a segment share similar properties. The contour's vertices p_n are visualised by lines perpendicular to the contour. Thick lines illustrate the segment boundaries.

While OLSTAD's contour description is solely grammatical, the fuzzy snake represents object boundaries in a form closer to natural language (cf. section 4.2.2). Also unlike OLSTAD's approach, a (fuzzy) segment length can be specified for each contour segment (cf. section 4.2.3), with the fuzzy contour describing the segment properties through fuzzy rule bases rather than algebraic functions (cf. section 4.2.4). The detailed realisation of these features, original to the fuzzy active contour, are described in section 4.3.

4.2.2 Contour Description

The grammatical approach described in the previous section extends the representational power of the traditional active contour. It allows for a more detailed exploitation of *a priori* knowledge by using different energy functions for subsequent segments of an object's boundary. This conversely may considerably increase the number of energy functions, their parameters and the weighting parameters between them, and often these numerical parameters are obscure and their refinement time consuming.

This is especially true in situations where the *a priori* knowledge is imprecise or uncertain, or where it is only available in the form of verbal expressions. It then becomes much more desirable to use a more intuitive, verbal contour description. The fuzzy snake allows for a description that is close to a natural language description of an object's boundary.

To illustrate this approach consider the local shape of the example object in Figure 4.5 again. The boundary may be decomposed into four segments. With OLSTAD's model, a grammatical expression to describe this contour could be

$$\alpha = \text{a*b*c*b*}$$

where a represents an energy function favouring local angles of 160°, while b and c favour angles of 180° and 110° respectively. A much more intuitive description would be

"A medium length arc, bending right, followed by a medium straight line,

a right bending corner and a short straight line."

This description can be formalised, where each contour segment is characterised by both its length and a property label. Table 4.1 provides the resulting linguistic contour description.

D = *medium* right arc, *medium* straight line,

very short right corner, *short* straight line

Table 4.1: Specification of fuzzy contour description in terms of each segment's length and property.

Syntactically a fuzzy contour description D is a concatenation of segment descriptions d_z, where $z \in [0, Z[$ with Z being the number of boundary segments. Each d_z is decomposed into a fuzzy segment length l_z and a segment property z_z.

$$D = \bigcup_{z=0}^{Z-1} d_z = \bigcup_{z=0}^{Z-1} l_z z_z \tag{4.1}$$

The details of these major components of the fuzzy snake's contour model are presented in the following sections.

4.2.3 Fuzzy Segment Length

The first element of the fuzzy snake representation is the specification of a segment's length l_z. To integrate linguistic values such as *medium*, *short*, and *very short* into an algorithm, they are translated into a number of snake vertices. In this way the fuzzy snake approach permits the specification of a number of subsequent vertices that share a common energy function, a parameter constraint which dramatically reduces the search space. A crisp length specification however, would not be able to consider

uncertain information. A new method to specify the length of a snake segment by a fuzzy number is therefore presented.

With OLSTAD's grammatical representation snake segments normally are of arbitrary length, for example α = a*b*c*b*. Only a crisp length of each segment could be given by concatenating the appropriate number of terminals as in the expression α = aaaaaabbbbbbcbbb. The fuzzy snake introduces a different notation, where the length is expressed as a fuzzy number (cf. section 3.2.1.1).

In the first segment of the contour in Figure 4.5, for a fuzzy length of $l_0 = 6$ and a spread of $s = 2$, the segment specification (6a) denotes a segment consisting of between 4 and 8 vertices, sharing a common property defined by an energy function E_{ext}^a.

Using the analogy of a grammatical description, the fuzzy snake can now be specified as an expression, for example D=(6a)(6b)(1c)(3b).

The absolute number of vertices, however depends on the sampling distance and is not a direct measure for the *length* of a segment. While the actual fuzzy snake algorithm uses the absolute length (cf. section 4.3.2), the user-level contour description must allow for a length measure relative to the overall length of the contour:

$$l_z^* = \text{round}(l_z/N) \tag{4.2}$$

With the segment length expressed as a percentage, the above example can be written as D=(0.4a)(0.4b)(0.05c)(0.15b). These relative lengths can now be mapped to linguistic labels as shown in Table 4.2 to realise a contour description as proposed in Table 4.1.

4.2.4 Fuzzy Segment Properties

A linguistic representation of a segment's properties is the second kernel element of the fuzzy snake model. Property labels such as right arc, straight line, and right

Linguistic label l_z	Relative length l_z^*	Absolute fuzzy length l_z
very short	5%	1
short	15%	3
medium	40%	6

Table 4.2: Mapping of linguistic labels to fuzzy numbers. Example for the fuzzy contour description in Table 4.1 with $N = 16$.

corner may be used instead of terminal characters such as a,b, and c of the grammatical model. This however, is only a replacement of arbitrary labels. The fuzzy snake conversely, uses linguistic fuzzy rules to describe all the features a contour may exhibit.

For this objective, a linguistic variable is created for each feature. To describe a local shape for example, the local angle at each vertex is measured and mapped to a linguistic variable *curvature*. Fuzzy sets are created and linguistic values assigned to characterise the curvature as for example *acute right* or *flat*. Taken together with an output variable *quality*, it is then possible to describe each segment's curvature by a fuzzy rule base. An example is given in Table 4.3.

The calculation of the active contour's external energy is performed by a fuzzy inference, the defuzzified output of which is a crisp quality measure for each contour candidate. The details of this approach are presented in section 4.3.3.5.

4.3 Realisation of the Fuzzy Active Contour

4.3.1 Processing a Multi-Segment Active Contour

Multiple external energy functions imply that a number of different calculations have to be performed for each vertex. Since the segment length is now variable, several energy functions may be calculated simultaneously for vertices either on or near a segment boundary. When back-tracking takes place in the dynamic programming optimisation,

Segment property label	Shortcut label	Fuzzy rule base
right arc	a	IF *curvature* medium right
		THEN *quality* very good
straight line	b	IF *curvature* flat left
		OR *curvature* flat right
		THEN *quality* very good
right corner	c	IF *curvature* strong right
		THEN *quality* very good

Table 4.3: Property labels defined by linguistic rules. Example for the fuzzy contour description in Table 4.1

only one energy function for each vertex succeeds. The selected function thereby determines the final *state* of that vertex.

With the fuzzy active contour each segment's length is specified as a fuzzy number. The realisation of this approach requires two separate operations to be performed:

1. The mean and the spread of l_z defines the possible states each vertex can be in. It hence defines which external energy functions have to be calculated for each vertex. With the fuzzy active contour an external energy function is determined by its specific rule base. Consequently when calculating the external energy of a particular candidate, the appropriate rule base has to be choosen. This operation is performed in module A6 (cf. Figure 4.6) and described in the following subsections.

2. The fuzzy membership degrees of l_z are taken into account by a modification of the dynamic programming optimisation algorithm and is described in section 4.3.2.
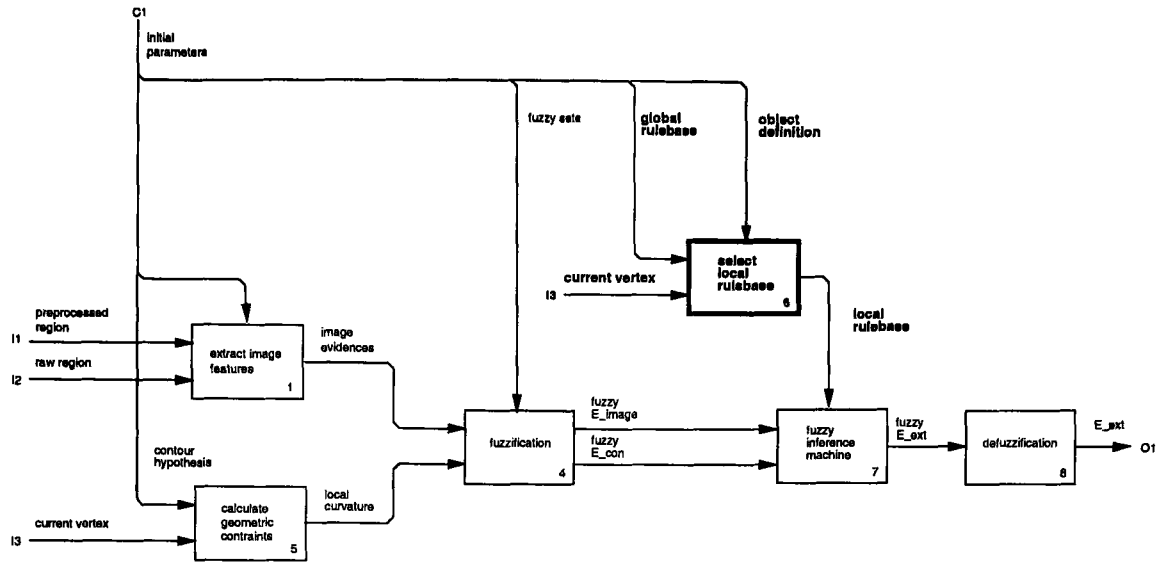
Figure 4.6: Principle structure of external energy calculation for the fuzzy active contour. Marked is the module to select the local rule base (A6).

4.3.1.1 Unspecified Segment Length

For a contour specification where the segment length is not specified as for example with OLSTAD's grammatical expression (cf. section 3.1.2.3) many alternative contours are valid. Consequently for most vertices all the properties' energy functions used in an expression α have to be calculated. This can be visualised in a function table. For the example used in the previous sections the expression α = a*b*c*b* leads to the function table shown in Table 4.4.

A resulting contour can comply with any pattern that can be generated through α, such as aaaaaaaaaaaabcb. An alternative formalism showing this are state graphs of finite state machines (FSMs) designed to generate or recognise all possible patterns for α^2 (cf. Figure 4.7).

---

[2]The state graphs used in this thesis follow the MOORE definition. A double circle denotes the start state. The state name corresponds to the name of the terminal characters used in the regular expression. A state transition occurs when the state machine is in a state and sees the character given at the arrow pointing to the target state.

Vertex	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	a	a	a	a	a	a	a	a	a	a	a	a	a			
Proper-		b	b	b	b	b	b	b	b	b	b	b	b	b		
ties			c	c	c	c	c	c	c	c	c	c	c	c	c	
				b	b	b	b	b	b	b	b	b	b	b	b	b

Table 4.4: Energy functions which have to be considered in the Figure 4.5 example, when the segment length is not specified.
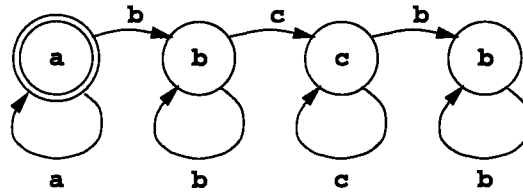

Figure 4.7: State graph for a finite state machine to generate the patterns defined by $\alpha = $ a*b*c*b*.

In theory the patterns generated by this FSM are of infinite length. In practice contours have a finite number of vertices (N), so that a FSM will perform only a finite number of state transitions. Consequently an FSM can be resolved, or expanded, using a state graph with no backward referencing links as shown in Figure 4.8.
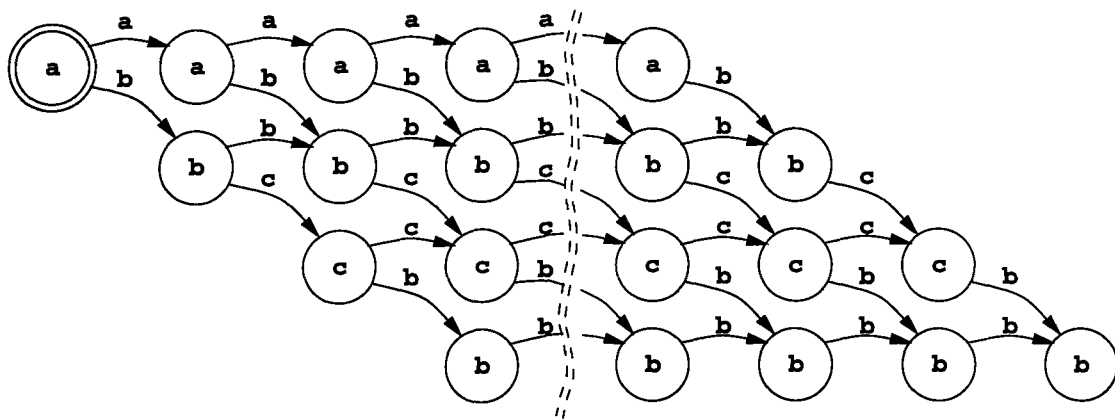

Figure 4.8: Expanded state graph for $\alpha = $ a*b*c*b*.

This graph however was simplified to show the formal background of the function table, their relation now being obvious: a column of states relates to the possible states a vertex may be in. A correct expansion has to assign different state names each time a transition is made, as the states do not reflect a vertex' state but the state of the whole system. The system state here relates to the state of the active contour.
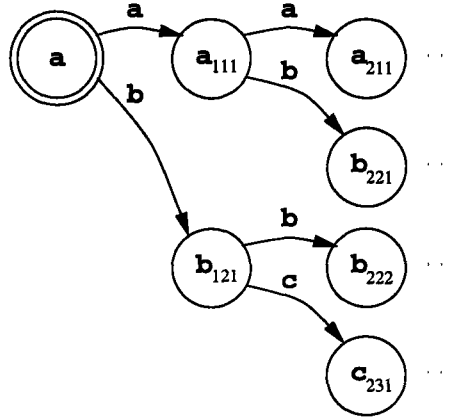

Figure 4.9: First nodes of a completely resolved state graph.

The first nodes of a completely resolved FSM are shown in Figure 4.9. Here the b state in the third column is split into two alternative states b_{221} and b_{222}. This demonstrates an important aspect of active contour optimisation: *all* possible paths have to be evaluated before the optimal pattern can be found, each vertex' final state is known and hence the contour is detected.

In the above example the overall energy of the first three vertices may be different when either the energy function sequences aab or abb are evaluated. Hence the active contour will either be in state b_{221} or b_{222}.

4.3.1.2 Specified Segment Length

With the absolute fuzzy segment length l_z (cf. section 4.2.3) the function table for the example contour description $D = l_0z_0\, l_1z_1\, l_2z_2\, l_3z_3 = (6a)\,(6b)\,(1c)\,(3b)$ is shown in Table 4.5.

Vertex	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Proper-ties	a	a	a	a	a	a	a	a								
					b	b	b	b	b	b	b	b	b	b	b	
									c	c	c	c	c	c	c	
										b	b	b	b	b	b	b

Table 4.5: Energy functions which have to be considered in the Figure 4.5 example, when the segment length is specified as a fuzzy number.

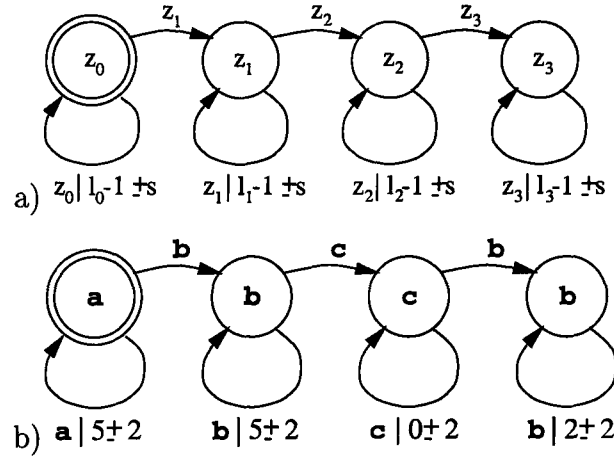The related state graph is given in Figure 4.10.

Figure 4.10: State graph for a contour description based on fuzzy active contours. a) general form for $D = l_0 z_0 \, l_1 z_1 \, l_2 z_2 \, l_3 z_3$, b) example for $D = (6\text{a})(6\text{b})(1\text{c})(3\text{b})$.

Unlike traditional state graphs the maximum number of state transitions is specified using the absolute fuzzy length l_z and its spread s. In practice some restrictions apply in extreme cases, such as that $l_z - 1 \pm s \geq 0$, of course.

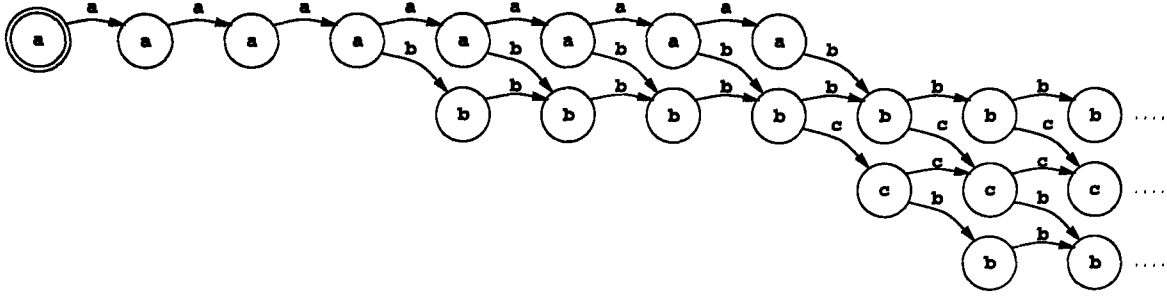The resolved and simplified state graph is given in Figure 4.11.

Figure 4.11: Expanded state graph for D = $(6a)(6b)(1c)(3b)$.

Incorporating the segment length reduces the computational complexity considerably. An example presented in [4] required only 2% of the number of states that were needed without the length constraint. It must be stated however that the actual computational complexity is very much application-specific and depends on l_z, s and the complexities $O(E^z_{ext})$, where $Z \in \{a, b, c, \ldots\}$.

With respect to module A6 of Figure 4.6 the global rule base is the complete set of rule bases defined for a given application (a, b and c in the example). The local rule base is selected by the described algorithm. The object definition is D.

4.3.2 Integration of a Fuzzy Segment Length

The actual implementation of the fuzzy part of the fuzzy segment length is achieved by an extension to the expression in line 8 of Table 3.2. The new expression is rewritten in line 3 of Table 4.6. The length is regarded as an additional constraint, where $l_z(n-1, k)$ is analogous to $S_z(n-1, k)$ and denotes the number of preceding vertices which would fall into state z if candidate $c_{n,m}$ was selected. Consequently, candidates which assist in constructing a chain of the specified length are favoured.

As the dynamic programming algorithm integrates (accumulates) local energy values, the length constraint must be based on the derivative (difference function) of the fuzzy number's membership function, shown in Figure 4.12.

Line 4 of Table 4.6 updates the length information. T_z in line 5 points to that predecessor of $c_{n,m}$ which would be optimal if the final state of vertex n was z.

1. **for** $n = 2 \ldots N - 1$

2. **for** $m = 0 \ldots M - 1$

3. $S_z(n, m) = \min_k \big[w_{int} E_{int}(c_{n,m}) + w_{ext} E^z_{ext}(c_{n,m}) + S_z(n - 1, k) + 1 - \Delta\mu(l_z(n - 1, k)) \big]$

4. $l_z(n, m) = l_z(n - 1, k^{min}) + 1$

5. $T_z(n, m) = k^{min}$
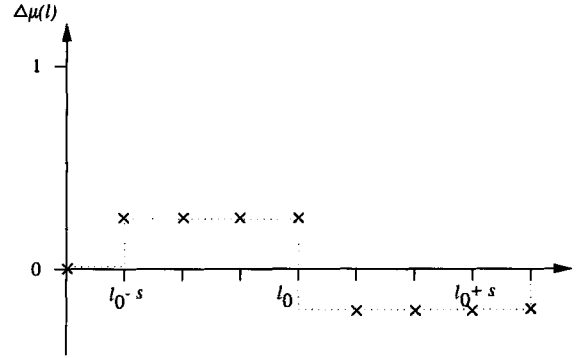
Table 4.6: Extended dynamic programming algorithm.

Figure 4.12: $\Delta\mu(l)$, as used in the length constraint of the extended dynamic programming algorithm.

4.3.3 Integration of Fuzzy Segment Properties

Snake applications are often insufficient. This is not a principle problem of active contours. The reason is rather that it is necessary to fully exploit the potential of the approach by formulating the energy function according to the characteristics of image and object. In the proposed system, this has been achieved by:

- introducing multiple image evidences from several sources to form E_{image},

- introducing an object specific constraint (E_{con}), that restricts the possible shapes of the active contour beyond a mere, general demand for smoothness.

Sections 4.3.3.1 and 4.3.3.2 introduce this framework and formulate appropriate energy functions that were applied successfully to medical image sequences (cf. section 5.2) using a traditional active contour. It is also shown where these functions are integrated into the fuzzy snake. Section 4.3.3.3 demonstrates how such energy functions are fuzzified in order to be processed in the fuzzy domain of the fuzzy active contour. As the fuzzy snake is generally capable of modelling arbitrary properties the described functions represent examples.

4.3.3.1 Algebraic Integration of Multiple Image Features

To exploit the benefits of the many existing image processing operators and at the same time compensate for their individual deficiencies, a structure is introduced that combines several edge-, region-, and motion-based low-level operators. With the fuzzy snake these operators are integrated according to Figure 4.13.
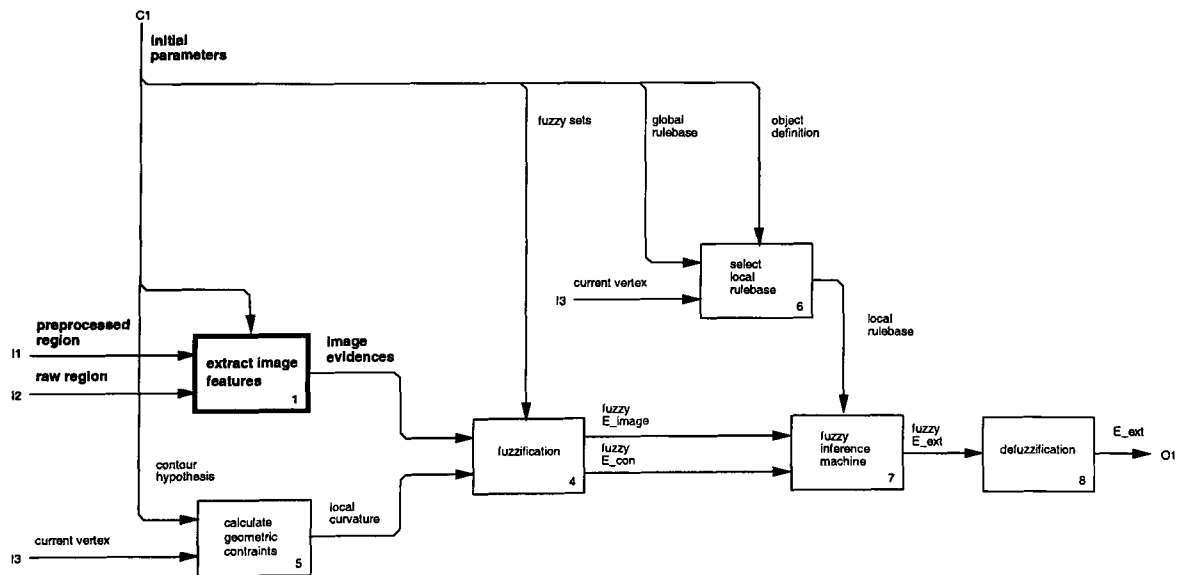

Figure 4.13: Principle structure of external energy calculation for the fuzzy active contour. Marked is the module to obtain edge-, region-, and motion-based image features (A1).

With the traditional active contour which is applied in chapter 5, the operators are combined in a computationally efficient manner by a normalised weighted sum (eq. 4.3).

$$E_{image} = 1 - e_{image} \qquad (4.3)$$

$$e_{image}(c_{n,m}) = w_{edge}||e_{edge}(c_{n,m})|| + w_{region}||e_{region}(c_{n,m})|| + w_{motion}||e_{motion}(c_{n,m})|| \qquad (4.4)$$

The result of each operator, normalised to the range $[0,1]$ through the norm $|| \; ||$, is regarded as an indication of to what degree a certain pixel c may belong to the object's contour. This local feature value is therefore referred to as *image evidence e*. It is also possible to incorporate operators that are able to deliberately detect structures which do *not* belong to the desired contour.

Each operator may wrongly lead to low evidence values for contour pixels and/or high values for non-contour pixels. The weighted combination of all evidences for a certain pixel compensates for the effect and leads to an attenuation of evidences of most contour pixels.

The individual evidence functions are very much application dependent. The following definitions in equations (4.5) to (4.7) have successfully been applied to the applications presented in this thesis.

The edge-based evidence is given by

$$e_{edge}(c_{n,m}) = \nabla_{\perp}(C_n, c_{n,m})\delta \qquad (4.5)$$

where $\nabla_{\perp}$ is the gradient normal to the initial contour hypothesis. The factor δ specifies either a rising ($\delta = 1$) or falling edge ($\delta = -1$) respectively.

The region-based evidence given in equation (4.6) weakens the influence of occluding objects with grey levels below a prescribed threshold θ. Pixels within these objects are

not influenced.

$$e_{region}(\mathbf{c}_{n,m}) = -|\nabla(\Theta, \mathbf{c}_{n,m})| \qquad \Theta(x_i, y_j) = \begin{cases} 1 & : \quad R(x_i, y_j) \le \theta \\ 0 & : \quad R(x_i, y_j) > \theta \end{cases} \qquad (4.6)$$

Here $R(x_i, y_j)$ denotes a pixel from the region of interest $\mathbf{R}$ around the contour hypothesis (cf. Figure 3.1). θ depends on *a priori* knowledge.

To exploit information from motion in the image sequence the third evidence value in equation (4.7) is applied.

$$e_{motion}(\mathbf{c}_{n,m}) = \varepsilon_{med}(\mathbf{c}_{n,m}), \qquad \varepsilon_{med} = \sqrt{\nabla(|\mathbf{R}_{t-d_1} - \mathbf{R}_t|)}\sqrt{\nabla(|\mathbf{R}_t - \mathbf{R}_{t+d_2}|)} \qquad (4.7)$$

This moving edge detector which is indicated by the subscript *med* [7], gives high evidence values for moving edges by multiplying the gradient of two difference images [152, 153]. The operation is performed on the same region of interest $\mathbf{R}_i$ extracted from images at instances $i = t$, $i = t - d_1$, and $i = t + d_2$, where d_1 and d_2 are constants.

4.3.3.2   Algebraic Integration of Object Features

For the traditional active contour a geometrical constraint presented in [1] is applied, which explicitly introduces knowledge on the expected shape of the contour and which can be easily incorporated into the dynamic programming algorithm. Equation (4.8) favours convex or concave contours by weighting the angle γ which is measured in the open polygon consisting of the current candidate vertex $\mathbf{c}_{n,m}$, its possible predecessor $\mathbf{c}_{n-1,k}$, and the optimal predecessor of the latter, $\mathbf{c}_{n-2,T(n-1,k)}$ (cf. eq. (4.9)). By always measuring γ in the same direction, setting the possible range $[\gamma^{min}, \gamma^{max}]$, the expected angle $\bar{\gamma}$ and the standard deviation σ, the resulting contour can be forced to bend in a desired direction (cf. eq. (4.10)).

$$E_{con}(\mathbf{c}_{n,m}) = ||\nu(\gamma(\mathbf{c}_{n,m}))|| \qquad (4.8)$$

$$\gamma(\mathbf{c}_{n,m}) = \angle(\mathbf{c}_{n-2,T(n-1,k)}, \mathbf{c}_{n-1,k}, \mathbf{c}_{n,m}) \tag{4.9}$$

$$\nu(\gamma) = \begin{cases} 0 & : \ \gamma^{\min} > \gamma > \gamma^{\max} \\ \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{\gamma-\bar{\gamma}}{\sigma}\right)^2} & : \ \gamma^{\min} \leq \gamma \leq \gamma^{\max} \end{cases} \tag{4.10}$$

The use of a Gaussian weighting function (again normalised to the range $[0, 1]$ through a normalisation function $\| \ \|$) allows exploitation of uncertain knowledge concerning the actual shape of the object. For instance if σ is small then the tolerance range around $\bar{\gamma}$ where E_{con} gives a good assessment is narrower than for higher values of σ. In theory, the weighting function $\nu(\gamma)$ applied in equation 4.10 could be replaced by any mathematical distribution which exhibited the appropriate properties.

With the fuzzy snake the weighting function is not used to realise a shape constraint. Only the shape feature γ is calculated in module A5 (cf. Figure 4.14).
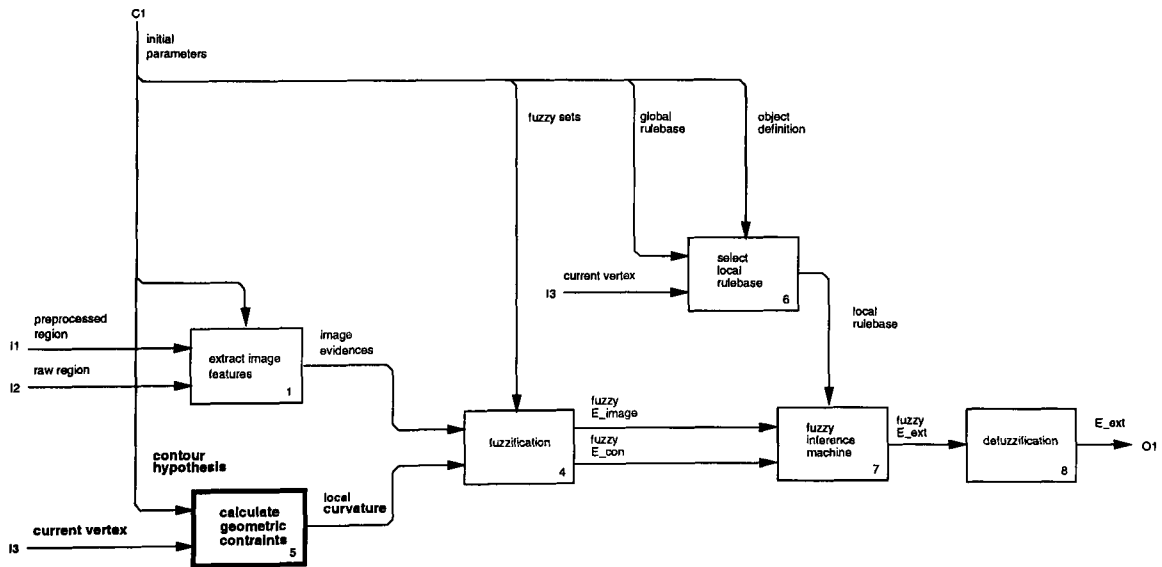

Figure 4.14: Principle structure of external energy calculation for the fuzzy active contour. Marked is the module calculating a shape feature (A5).

4.3.3.3 Fuzzification

With the fuzzy snake, the components of the external energy function (in equations (4.3) and (4.8)) are separately represented by linguistic variables and fuzzy sets. The transformation from the numeric to the fuzzy domain (the fuzzification) is performed in module A4 of the fuzzy active contour's external energy calculation (cf. Figure 4.15).

With E_{image} the calculation of image evidences (equation (4.5) to (4.7)) is retained, but they become linguistic variables.
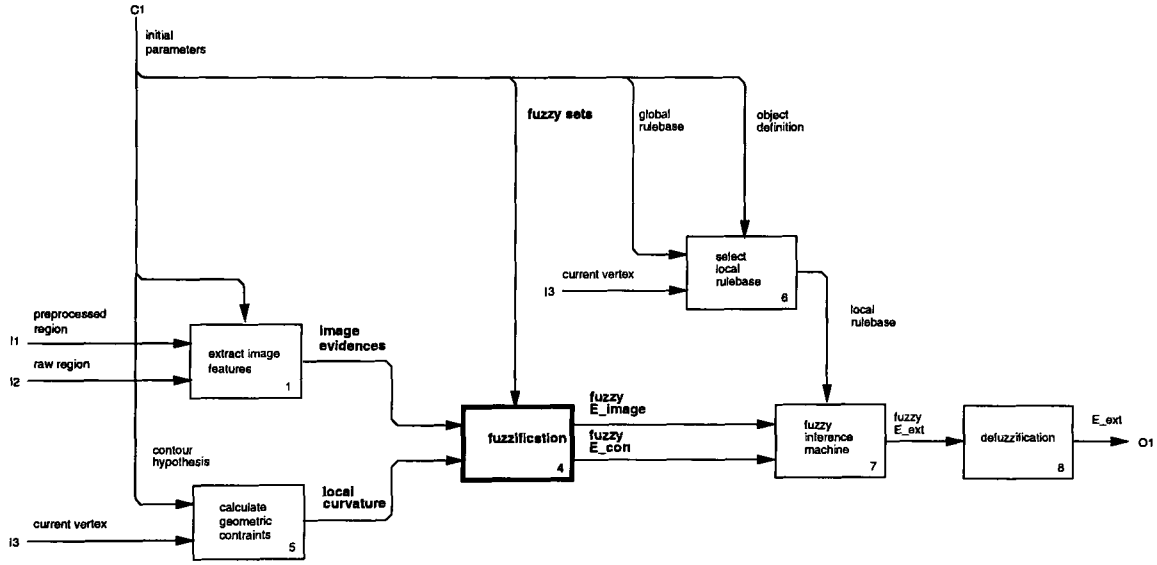

Figure 4.15: Principle structure of external energy calculation for the fuzzy active contour. Marked is the fuzzification module (A4).

The following example illustrates how an evidence function can be extended by an intuitive interface using elements of fuzzy logic. For e_{edge}, a linguistic variable $edge$ is created. Fuzzy sets with linguistic values, for instance falling very strong, ..., rising very strong are defined to cover the value range of equation (4.5).[3] In this example, the values of the term $\nabla_\perp(C_n, c_{n,m})$ are mapped to adjectives which cover the range from

---

[3]Note for simplicity, that the adjectives are regarded as a part of the primary term of the linguistic value, rather than as a linguistic hedge with an associated operator in the sense of [16].

very weak to very strong. The factor δ which specifies the edge direction, is accounted for by the adjectives rising and falling respectively.

The formal definition for this linguistic variable is

$$u_1 = e_{edge}, \qquad \mathcal{U}_1 = [-1, 1], \qquad x_1 = edge$$

$\mathcal{T}_1(edge)$ = { falling very strong, falling strong, falling medium, falling weak,

falling very weak, rising very weak, rising weak, rising medium,

rising strong, rising very strong}

The definitions of the related membership functions are shown in Figure 4.16.
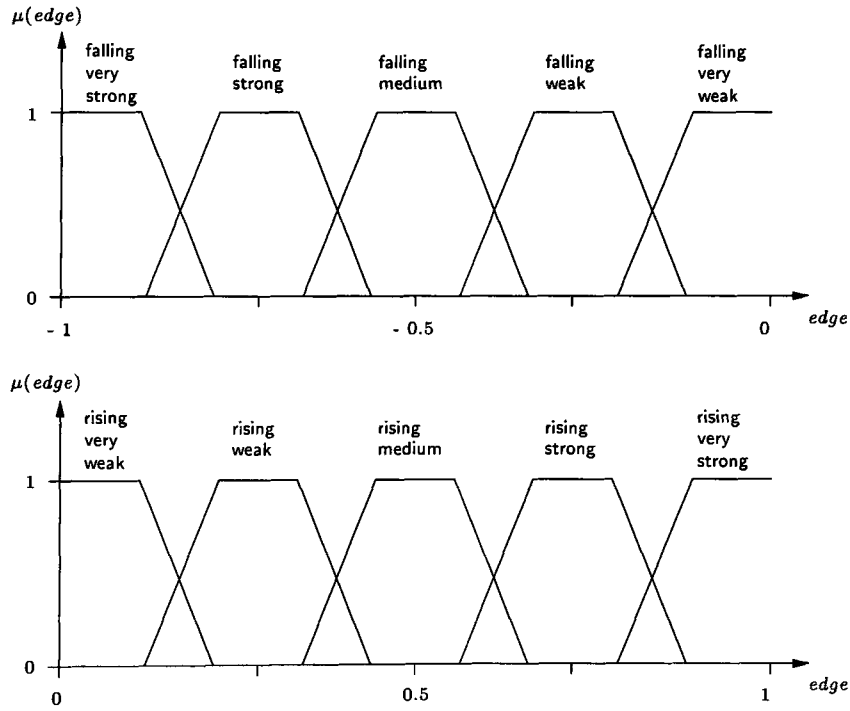

Figure 4.16: Fuzzy representation of edge-based image evidences.

In an analogous manner, the definition of linguistic variables and fuzzy sets for the region and motion-based evidences (equations (4.6) and (4.7) respectively) are similarly given.

$$u_2 = e_{region}, \qquad \mathcal{U}_2 = [-1, 1], \qquad x_2 = region$$

$\mathcal{T}_2(region) = \{$ negative very strong, negative strong, negative medium, negative weak,

                negative very weak, positive very weak, positive weak, positive medium,

                positive strong, positive very strong$\}$

$$u_3 = e_{motion}, \qquad \mathcal{U}_3 = [-1, 1], \qquad x_3 = motion$$

$\mathcal{T}_3(motion) = \{$ negative very strong, negative strong, negative medium, negative weak,

                negative very weak, positive very weak, positive weak, positive medium,

                positive strong, positive very strong$\}$

The definitions of the related membership functions are identical to those shown in Figure 4.16.

To illustrate how a constraint energy function is integrated into the fuzzy snake, the fuzzification of equation (4.8) is demonstrated. The geometrical constraint becomes a linguistic variable *curvature* which is calculated using equation (4.9) and represents the actual measure for the constraint.

The weighting function ν in equation (4.10) however, is now replaced by a number of fuzzy sets. The membership functions of the fuzzy sets acute to flat have been determined empirically to relate to the human perception of the different degrees of curvature. Prototype polygons such as those shown in Figure 4.17 were classified by a number of test persons. The statistical distribution of these classifications are the basis for the membership functions of the linguistic variable *curvature* and are shown in Figure 4.18.
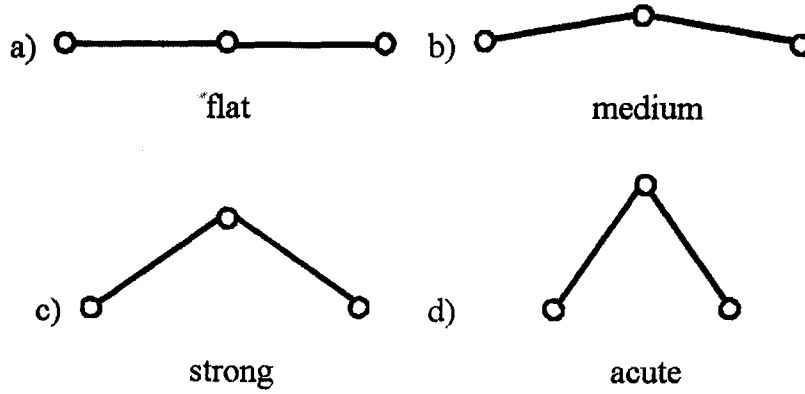
Figure 4.17: Examples for the 3-vertex polygons that were used in empirically determining the membership functions
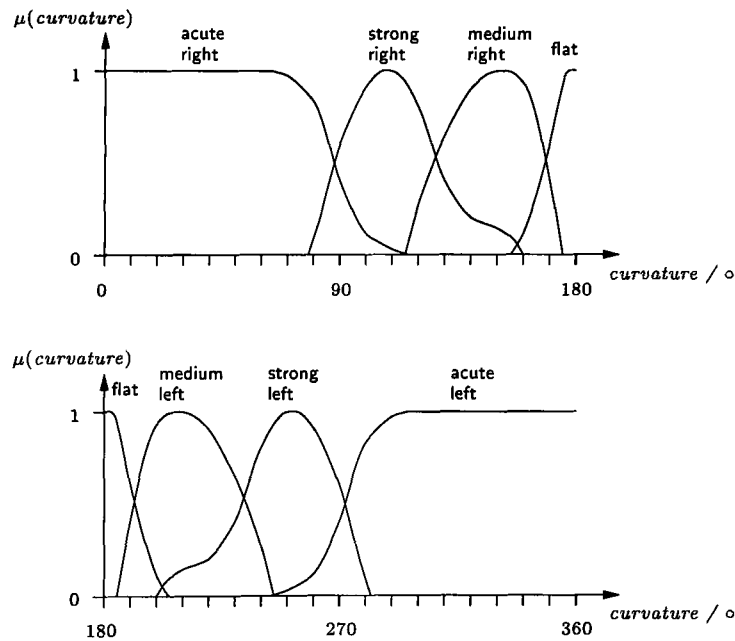

Figure 4.18: Fuzzy representation of curvature constraint.

Due to the fuzzification of the curvature measure the desired curvature no longer has to be expressed through the parameters $\bar{\gamma}$, γ^{min}, γ^{max} and σ, but instead more intuitive linguistic values may be used.

The formal definition of *curvature* is

$$u_4 = \gamma, \qquad \mathcal{U}_4 = [0, 360], \qquad x_4 = curvature$$

$$\mathcal{T}_4(curvature) \;=\; \{ \quad \text{acute right, strong right, medium right, flat,}$$
$$\text{medium left, strong left, acute left}\}$$

Many other alternative constraints may also be included, for example the orientation of a contour segment (by using the angle of the major axis of a segment) or the relative position of different contour segments.

The entire membership degrees $\mu(edge)$, $\mu(region)$ and $\mu(motion)$ for a given candidate $c_{n,m}$ can be regarded as its fuzzy image energy,[4] and membership degrees for constraints such as $\mu(curvature)$ correspondingly as its fuzzy constraint energy.

To complete the fuzzification an output linguistic variable is defined. It allows for the definition of linguistic rules which specify the quality of a vertex:

$$E_{ext} = 1 - u_o, \qquad \mathcal{U}_o = [0, 1], \qquad z_o = quality$$

$$\mathcal{T}_o(quality) = \{\text{very bad, bad, medium, good, very good}\}$$

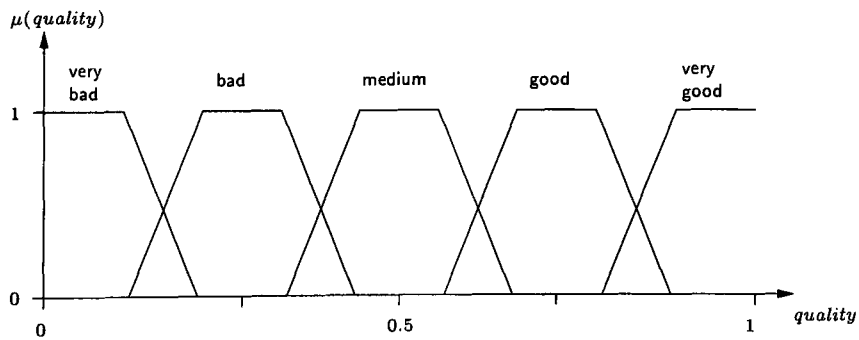The membership functions are defined as shown in Figure 4.19.

Figure 4.19: Output variable of the fuzzy system.

---

[4]It should be recalled that high membership values relate to low energy values (see definition of equation (4.3)).

4.3.3.4 Linguistic Rules

Algebraic equations (4.3) and (4.8) can now be replaced by a fuzzy inference process using formal linguistic rules instead of trial-and-error weights.

As previously described, each energy function for a boundary segment exhibiting constant properties consists of a rule base, which inputs a number of different evidences or features from the image, as well as constraints on for example, the geometry of an object. The output from the inference process using this rule base, is a *quality* measure (cf. Figure 4.19), describing the compliance with the rule base for each vertex. Normally rules will have the conclusion *quality* very good to describe a known desired result, while negative linguistic values for *quality* may be used for those properties which a contour segment must *not* exhibit.

Considering the edge and curvature properties of the example shown in Figure 4.5, the contour can now be described by the four rule bases labelled a, b, c and d respectively in Table 4.7. In this simple example each rule base consists of only one rule R_1.

a:

R_1 : IF *edge* falling very weak

AND *curvature* medium right

THEN *quality* very good

b:

R_1 : IF *edge* falling weak

AND (*curvature* flat left

OR *curvature* flat right)

THEN *quality* very good

c:

R_1 : IF *edge* falling medium

AND *curvature* strong right

THEN *quality* very good

d:

R_1 : IF *edge* falling medium

AND (*curvature* flat left

OR *curvature* flat right)

THEN *quality* very good

Table 4.7: Rule bases for the example in Figure reffig:ex3.

4.3.3.5 Fuzzy Inference

For each calculation of an external energy one rule base is inferenced according to the procedures described in section 3.2.1.5.

4.3.3.6 Defuzzification

The defuzzification procedure converts the result of the inference process which is a fuzzy quality measure, into a crisp value $E_{ext}^z(c_{n,m})$. Here the centre-of-gravity method is used for defuzzification.

4.4 Calculation Example

This section summarises the calculations performed to obtain the external energy function E_{ext} of a fuzzy snake and gives example values. The calculations are performed for one candidate vertex $c_{n,m}$.

Evidence calculation

$$e_{edge}(c_{n,m}) = ||\nabla_\perp(C_n, c_{n,m})||_{-1..1} = -0.65 \qquad (4.11)$$

$$e_{region}(c_{n,m}) = ||\nabla(\Theta, c_{n,m})||_{-1..1} = -0.1 \qquad (4.12)$$

$$e_{motion}(c_{n,m}) = ||\varepsilon_{med}(c_{n,m})||_{-1..1} = 0.8 \qquad (4.13)$$

$$\gamma(c_{n,m}) = ||\angle(c_{n-2,T(n-1,k)}, c_{n-1,k}, c_{n,m})||_{0°..360°} = 120° \qquad (4.14)$$

Rule base selection

The actual selection of a rule base to be applied to one candidate is done within the overall algorithm, as described in section 4.3.1. Here we assume the following example rule base a:

R_1 : IF *motion* positive strong THEN *quality* very good

R_2 : IF *edge* falling weak THEN *quality* good

R_3 : IF *curvature* strong right THEN *quality* good

R_4 : IF *region* negative strong

    OR *region* negative very strong THEN *quality* bad

R_5 : IF *edge* falling medium

    OR *edge* falling strong THEN *quality* bad

With R_1 the motion evidence is modeled as a strong indicator for the desired object. R_2 and R_3 are less confident. R_4 models a distracting dark object, while R_5 models unwanted structures which have slightly stronger edges than the wanted object (cf. R_2).

Fuzzification

The fuzzy inference processes the rules in several steps. The first step is to fuzzify the input values, that is the evidence values obtained in equations (4.11) to (4.14).

Each value is input to its related linguistic variable *edge, region, motion* and *curvature* (cf. Fig. 4.16 and 4.18). For each of the fuzzy sets defined for a linguistic variable, the membership value is calculated. In practise, only fuzzy sets that are actually used in the rule base are calculated, for example:

for R_1 : $\mu^{motion}_{\text{positive strong}}(e_{motion}(\mathbf{c}_{n,m}))$ $= \mu^{motion}_{\text{positive strong}}(0.8)$ $= 0.9$

for R_2 : $\mu^{edge}_{\text{falling weak}}(e_{edge}(\mathbf{c}_{n,m}))$ $= \mu^{edge}_{\text{falling weak}}(-0.65)$ $= 0.8$

for R_3 : $\mu^{curvature}_{\text{strong right}}(\gamma(\mathbf{c}_{n,m}))$ $= \mu^{curvature}_{\text{strong right}}(120)$ $= 0.7$

for R_4 : $\mu^{region}_{\text{negative strong}}(e_{region}(\mathbf{c}_{n,m}))$ $= \mu^{region}_{\text{negative strong}}(-0.1)$ $= 0.0$

$\mu^{region}_{\text{negative very strong}}(e_{region}(\mathbf{c}_{n,m}))= \mu^{region}_{\text{negative very strong}}(-0.1) = 0.0$

for R_5 : $\mu^{edge}_{\text{falling medium}}(e_{edge}(\mathbf{c}_{n,m}))$ $= \mu^{edge}_{\text{falling medium}}(-0.65)$ $= 0.2$

$\mu^{edge}_{\text{falling strong}}(e_{edge}(\mathbf{c}_{n,m}))$ $= \mu^{edge}_{\text{falling strong}}(-0.65)$ $= 0.0$

Aggregation

The next step is to obtain truth values v for the premises P. Where rules consist of more than one premise an aggregation is required.

$$v_1 = \mu_{\text{positive strong}}^{motion}(e_{motion}(c_{n,m})) = 0.9$$

$$v_2 = \mu_{\text{falling weak}}^{edge}(e_{edge}(c_{n,m})) \qquad = 0.8$$

$$v_3 = \mu_{\text{strong right}}^{curvature}(\gamma(c_{n,m})) \qquad = 0.7$$

$$v_4 = v(P_{41} \text{ OR } P_{42})$$
$$= \max\left[\mu_{\text{negative strong}}^{region}(e_{region}(c_{n,m})), \mu_{\text{negative very strong}}^{region}(e_{region}(c_{n,m}))\right]$$
$$= \max[0.0, 0.0]$$
$$v_4 = 0.0$$

$$v_5 = v(P_{51} \text{ OR } P_{52})$$
$$= \max\left[\mu_{\text{falling medium}}^{edge}(e_{edge}(c_{n,m})), \mu_{\text{falling strong}}^{edge}(e_{edge}(c_{n,m}))\right]$$
$$= \max[0.2, 0.0]$$
$$v_5 = 0.2$$

Implication

Each rule's conclusion is now weighted by the related truth value, resulting in a fuzzy set Z for each rule.

$$Z_1 = \{(u, \min\left[v_1, \mu_{\text{very good}}^{quality}(u)\right])\} = \{(u, \min\left[0.9, \mu_{\text{very good}}^{quality}(u)\right])\}$$

$$Z_2 = \{(u, \min\left[v_2, \mu_{\text{good}}^{quality}(u)\right])\} \qquad = \{(u, \min\left[0.8, \mu_{\text{good}}^{quality}(u)\right])\}$$

$$Z_3 = \{(u, \min\left[v_3, \mu_{\text{good}}^{quality}(u)\right])\} \qquad = \{(u, \min\left[0.7, \mu_{\text{good}}^{quality}(u)\right])\}$$

$$Z_4 = \{(u, \min\left[v_4, \mu_{\text{bad}}^{quality}(u)\right])\} \qquad = \{(u, \min\left[0.0, \mu_{\text{bad}}^{quality}(u)\right])\}$$

$$Z_5 = \{(u, \min\left[v_5, \mu_{\text{bad}}^{quality}(u)\right])\} \qquad = \{(u, \min\left[0.2, \mu_{\text{bad}}^{quality}(u)\right])\}$$

Accumulation and Defuzzification

The accumulation of the above fuzzy sets results in a single fuzzy set, that reflects the fuzzy quality of the calculated candidate:

$$\mathcal{Z} = \{(u, \mu_{\mathcal{Z}}(u))\} = \{(u, \max_{k=1..5}[\mu_{\mathcal{Z}_k}(u)])\}$$

As the contour optimisation is performed in the crisp domain, a defuzzification is performed, to obtain a single crisp value from the resulting fuzy set:

$$u_o = \frac{\sum\limits_{u \in \mathcal{U}} u \, \mu_{\mathcal{Z}}(u)}{\sum\limits_{u \in \mathcal{U}} \mu_{\mathcal{Z}}(u)} = 0.8$$

High membership values correspond to a low contour energy, hence the crisp output value is inverted to obtain the external energy value for the calculated candidate:

$$E_{ext}^{a}(\mathbf{c}_{n,m}) = 1 - u_o = 0.2$$

4.5 Summary

This chapter has introduced the concepts and presented the theoretical basis of the fuzzy active contour.

This novel approach for the segmentation of deformable structures affords a seamless integration of traditional image processing operators, active contours and a fuzzy reasoning component. In summary, the main features of this concept are:

- Uncertainties immanent in linguistic *a priori* knowledge are exploited through an original contour description method.

- The description allows for the inclusion of both image- and object-related characteristics.

- Furthermore the linguistic description method eliminates the need for obscure numerical system parameters.

- Immanent uncertainties are propagated through most of the processing blocks of the system. This way, the necessary crisp segmentation decision is made at a very late stage, reducing segmentation errors considerably.

This chapter has also presented a computationally efficient realisation of the fuzzy active contour, based on a finite state machine and a dynamic programming optimisation algorithm. The principle concept however is independent of a particular implementation.

Chapter 5

Experimental Results

5.1 Application to Synthetic Images

5.1.1 Detection of Multi-Segment Object Contours

The refined model representation presented in this research improves contour detection in comparison to traditional active contours. To validate this characteristic synthetic images were created which allow for an experimental analysis. For clarity the only property considered was shape.

Without specific *a priori* knowledge an active contour cannot detect a specific object. To demonstrate this, Figure 5.1a shows two symmetrical objects with a constant local angle. An active contour with only a similarity constraint would normally detect the object most similar to the initial hypothesis. To detect the lower object, for example the hypothesis would have to be bend downwards as well.

In this example however, the hypothesis was placed on the axis of symmetry between both objects. As now both contours share the same similarity relative to the hypothesis, it depends on the implementation which contour is detected. Here the lower object was found (Figure 5.1b).

An active contour with a global shape constraint can represent more specific knowledge about the object. Hence it is possible to detect the lower object by setting the desired local angle $\bar{\gamma}$ to 170 degrees as shown in Figure 5.1c. Note that in this case the result is identical to that obtained by the similarity constraint, but that it was *desired* and does not depend on the particular implementation of the optimisation algorithm. Setting $\bar{\gamma}$ to 190 degrees detects the upper object (Figure 5.1d).
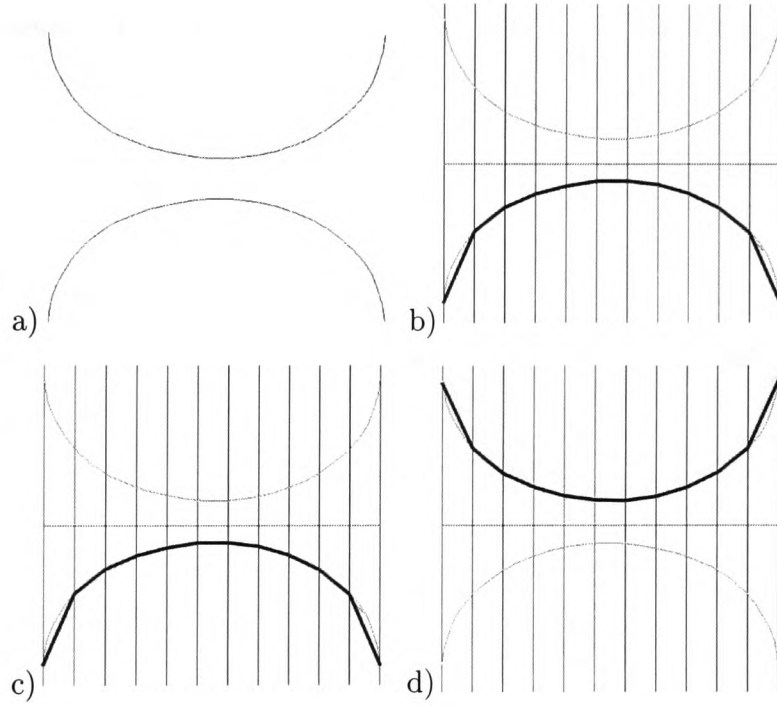

Figure 5.1: Detecting a specific object. a) two symmetric objects with constant local angles, b) implementation-dependent detection through a similarity constraint, c) and d) deliberate selection of either object through a global shape constraint. Figs. b, c and d also contain the initial contour hypothesis **P** (horizontal line) as well as the search lines $\mathbf{C}_i$ (vertical lines).

In the above example the global shape constraint was advantageous. To demonstrate its limits different local properties are introduced in parts of a contour (Figure 5.2a).

The similarity constraint chooses the constant contour (Figure 5.2b) because the hypothesis was constant. Unlike the example in Figure 5.1b this result is not implementation-

dependent. A hypothesis similar to the upper object had resulted in the detection of that contour.

The same constant contour was detected with a global shape constraint (Figure 5.2c). In these synthetic images without noise or distortions there are only a few alternative results possible so that the explicit formulation of a specific shape can be equivalent to the implicit model of the similarity constraint.

The traditional active contour cannot detect the upper inconstant object. This object requires a more detailed model which can be provided by the proposed fuzzy contour model.
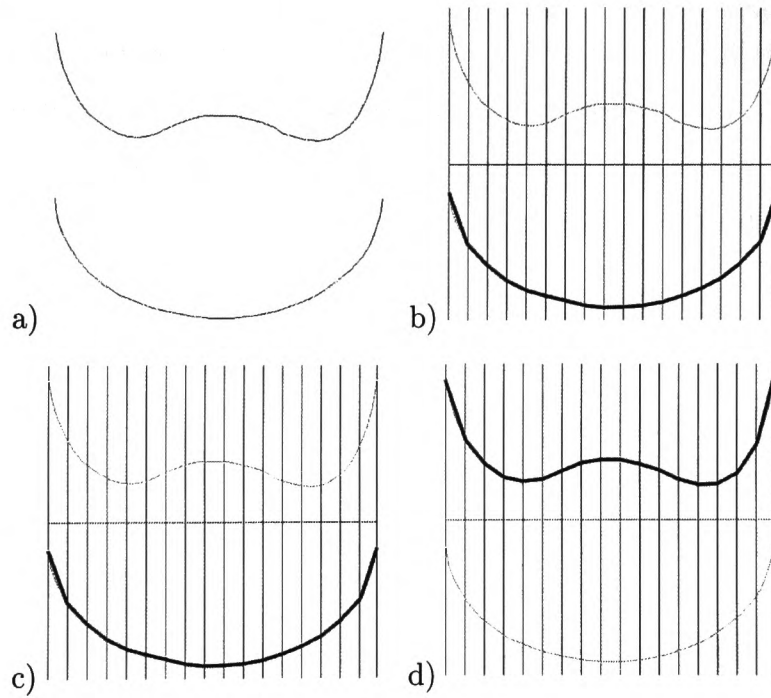
Figure 5.2: Detecting an object with inconstant local shape. a) two example objects, b) the similarity constraint detects the constant contour because it is relatively similar to the initial hypothesis, c) a global shape constraint detects the same contour because it matches absolutely, d) the more detailed model of the fuzzy snake can detect the inconstant object.

It was demonstrated that the proposed model is generally capable of selecting objects with a specific shape. An important aspect in a model which exploits uncertain

knowledge is its selectivity. The model should be flexible enough to cover uncertainty in both the object description and the image features. If this flexibility is too large however, the model cannot separate the desired object from unwanted distortions or a partially similar unwanted object. To demonstrate the sufficient ability of the fuzzy contour model to separate similar objects, two contours were superimposed in Figure 5.3a). Both contours have the same local angles, except for the middle section of one of the objects.

The similarity-sensitive active contour is not able to detect what a human observer would consider as one of two possible objects. The resulting contour merely is similar to the initial hypothesis (Figure 5.3b). The reason is that in this case the similarity constraint does not coincide with the general object model of the human visual system. We see two objects: a bowl and a flat inconstant contour. The fact that the bowl-shaped object was already presented in the above examples accentuates our perception of it (priming). Consequently the remaining contour is perceived as the second object. Such models are not included within the simple snake.
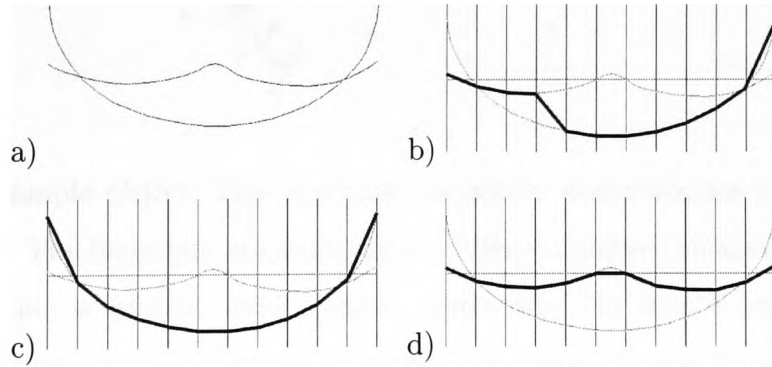
Figure 5.3: Distinguishing between perceptible objects having the same local shape.

Figure 5.3c demonstrates that a global constraint results in a detection of the constant object as was the case in the preceding examples. Although 10 out of 12 vertices have the same local angle the active contour finds the global optimum.

The inconstant object again requires a more detailed model which can be incorporated by the proposed fuzzy snake.

5.1.2 Fuzziness of Boundary Features

This section validates that fuzzy snakes are tolerant and flexible, with the limits of this flexibility explored.

The fuzzy snake model constrains the possible shapes the resulting contour might have, while concomitantly allowing for a certain deviation from an optimal prototype of the object that is to be detected. The variable segment length is one element which affords this flexibility. Another is the fuzziness in the description of desired boundary features. This is demonstrated on a synthetic object which is introduced in Figure 5.4.
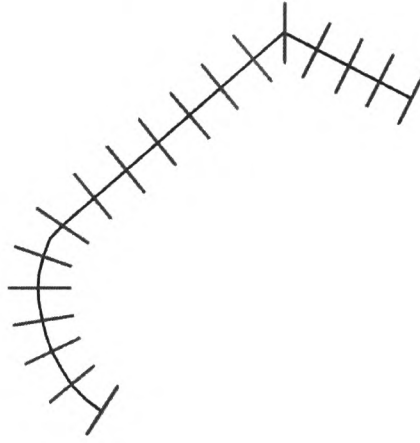
Figure 5.4: Example object. This synthetic boundary demonstrates a multi-segment object contour. The boundary segments have a different shape (local curvature), while all vertices within a segment share similar properties. The search lines used by the (fuzzy) active contour are visualised by grey lines perpendicular to the contour and intersecting the contour's vertices p_n.

Figure 5.5 presents an example similar to those used in the preceding section, which illustrates how by using a more sophisticated model detection of the example object can be successfully achieved.

To illustrate how the proposed fuzzy snake is capable of favouring desired features using the fuzzification described above, the following example visualises the membership degrees of the fuzzy sets over the linguistic variable *curvature*.
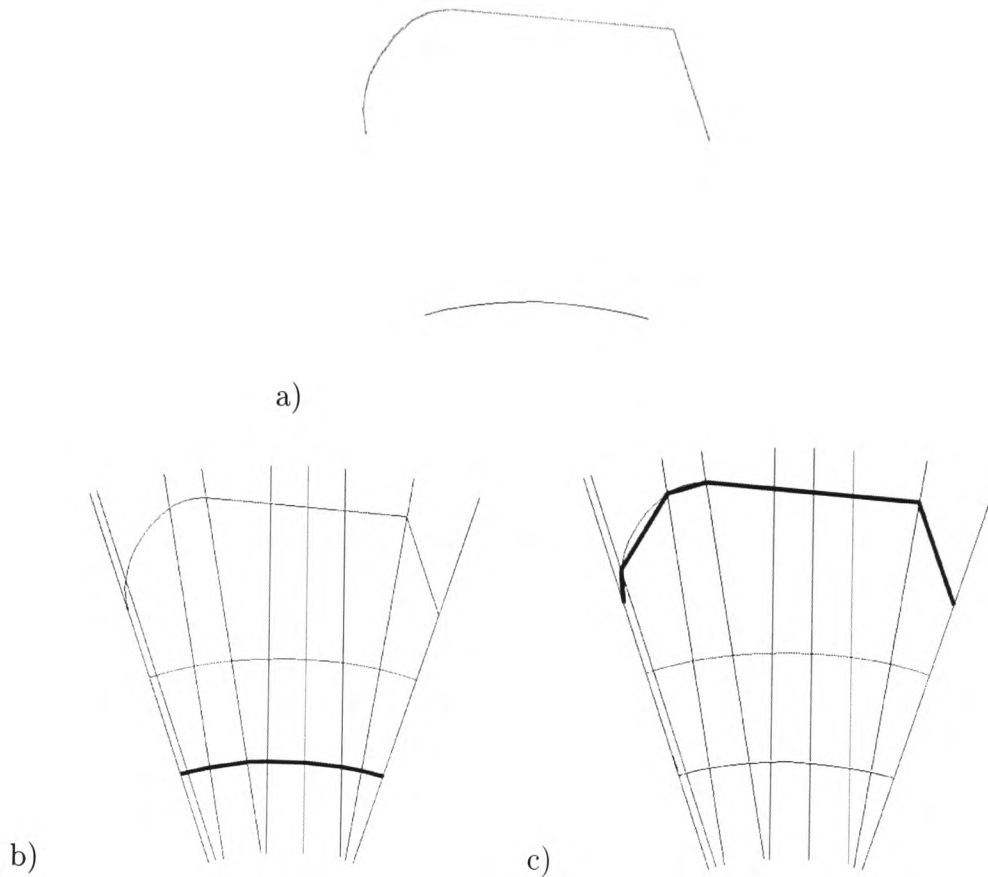
Figure 5.5: Two objects with different curvature properties (a). The lower constant contour can be detected by a traditional active contour (b), while the upper multi-segment contour requires a better model. Figs. b and c also contain the initial contour hypothesis $\mathbf{P}$ (upper arc) as well as the search lines $\mathbf{C}_i$ (straight lines).

For each vertex of the example contour in Figure 4.5, all membership degrees for the fuzzy sets in Figure 4.18 are calculated. The highest and second highest value (if any), for each vertex are recorded in the graph in Figure 5.6. The graph shows that the analysed contour is very similar to the description given in Tables 4.1 and 4.3, which demand curvatures of medium right, flat left OR flat right, strong right, flat left OR flat right.

These relations become more evident when a different visualisation is applied as demonstrated in Figure 5.7.
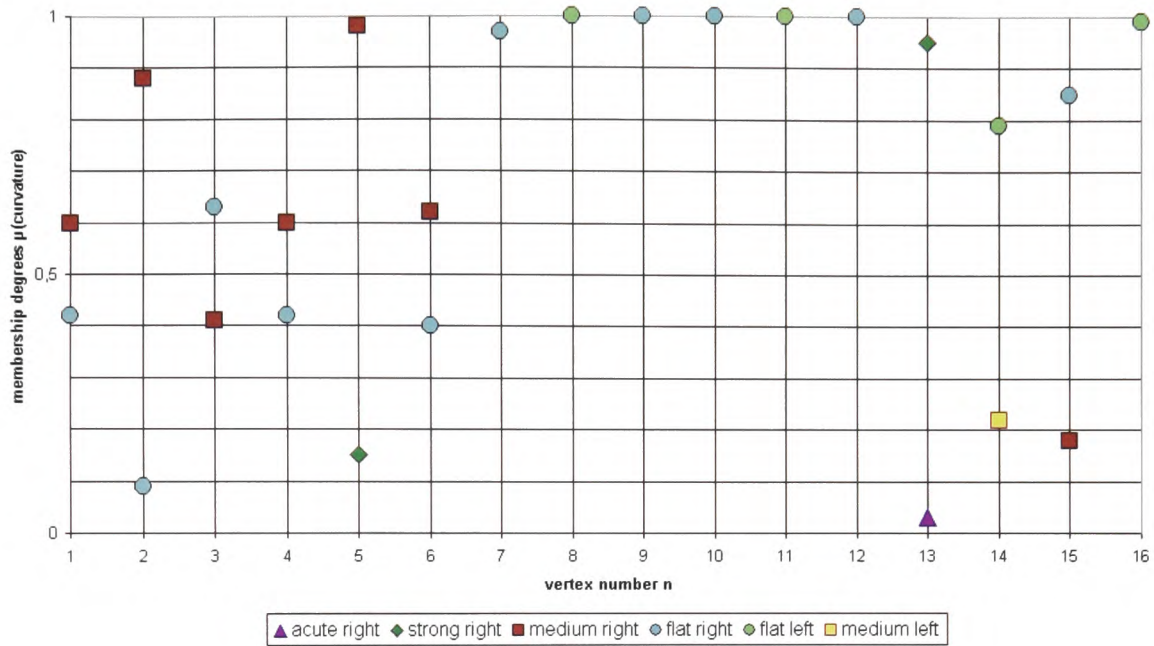
Figure 5.6: Significant membership degrees for each vertex of a contour.

Here the fuzzy sets are colour coded and superimposed as circles over the contour image. The circle diameter is proportional to the mean degree of membership within a consecutive sequence of equal fuzzy sets.

a)   b)
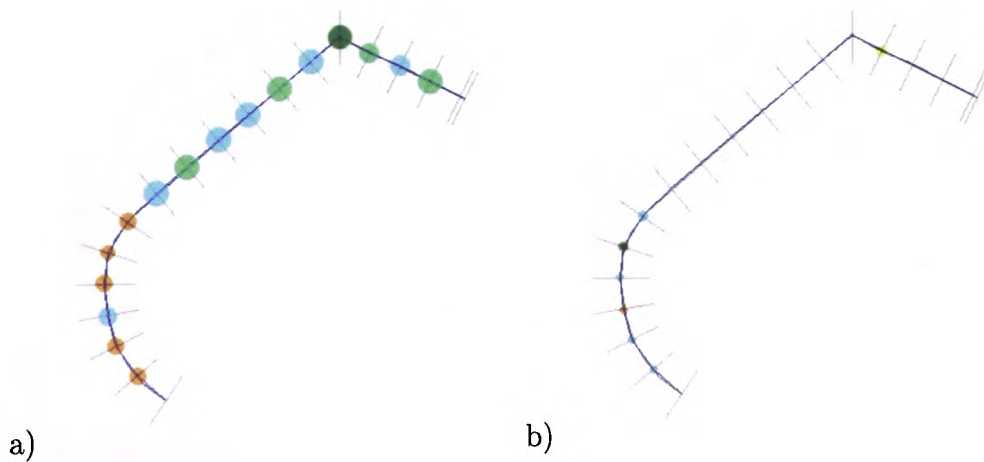
Figure 5.7: Visualisation of membership degrees for boundary features. The fuzzy sets are colour coded. a) maximum membership degrees of each vertex, b) second highest values.
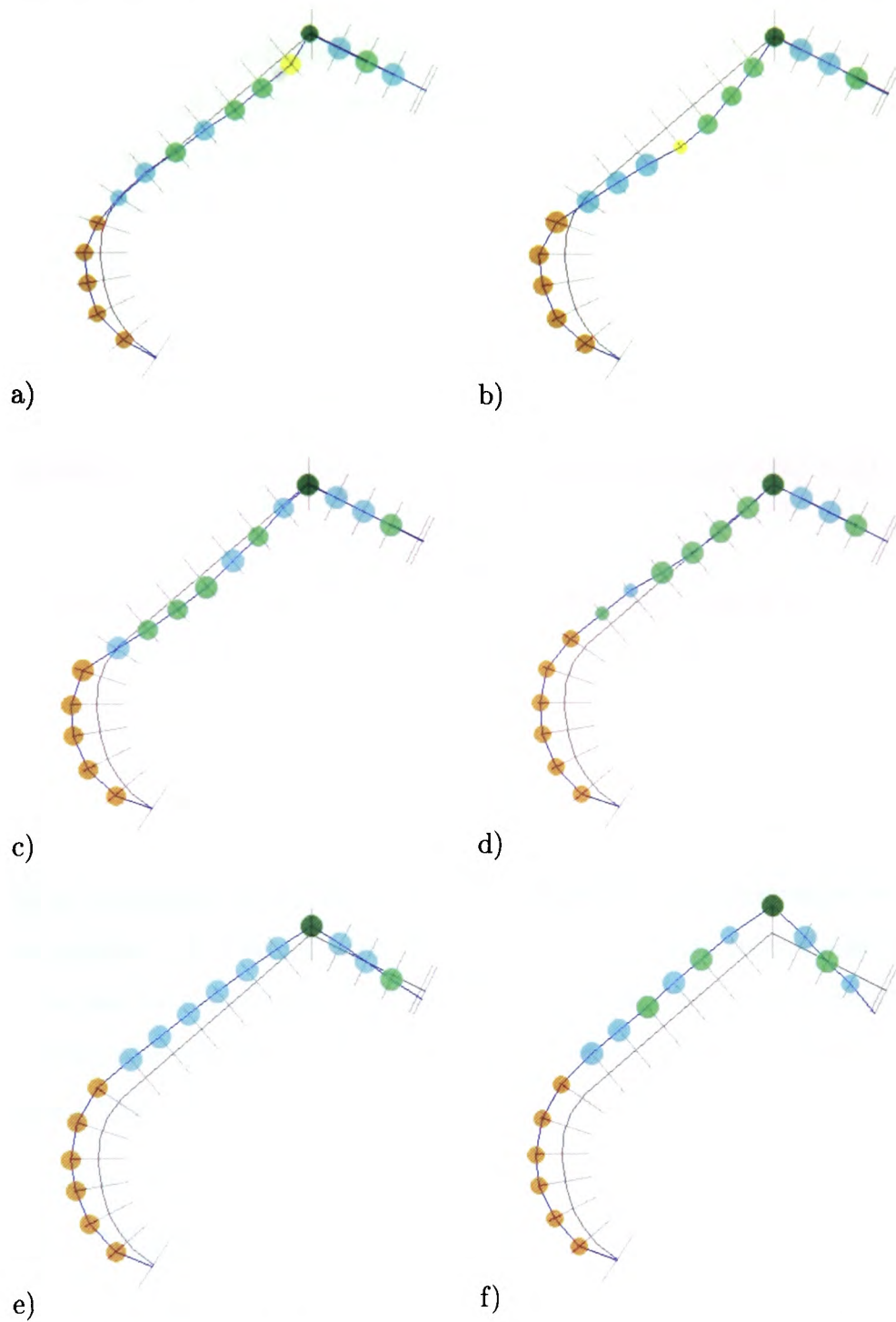
Figure 5.8: Demonstration of the shape feature's fuzziness through a visualisation of the mean maximum membership degrees for the curvature of different non-rigid deformations of a prototype.

As both the representation and detection of non-rigid objects are of particular interest, a sequence of deformations of the example contour is shown in Figure 5.8. This sequence demonstrates to what degree the fuzziness of a contour description favours certain deviations from the prototype while placing less emphasis on others. In Figure 5.8a) to c) a local distortion was introduced in the longer straight segment, bending the contour to the left. With the deformations 5.8a) and b) the local curvature at the distorted vertices results in a high value for $\mu_{\text{medium left}}(curvature)$, while the desired $\mu_{\text{flat left}}(curvature)$ is very small or zero (not shown). This means that the overall energy E_{snake} is significantly smaller for 5.8a) and b) than it is for 5.8c). In other words the deformation 5.8c) is more similar to the prototype than either 5.8a) or b).

5.1.3 Summary of Application to Synthetic Images

Traditional active contours are capable of representing object characteristics, such as shape, through a global similarity constraint. It was demonstrated that this approach is insufficient to detect complex contours.

While an extension of an active contour to represent local constraints would in theory be possible, this would require too many numerical parameters. Conversely the fuzzy active contour representation allows for the intuitive, linguistic representation of local characteristics, while being more compact than a mere numerical representation. It was shown, that contours with locally inconstant characteristics can be detected.

Furthermore the fuzzy active contour's selectivity in the presence of local deformations was explored. It was shown, that the linguistic terms used in a fuzzy contour description are in fact related to perceptible changes in an object's shape.

Synthetic images were also used in [4] to evaluate the fuzzy active contour's properties.

5.2 Application to Medical Images

Each of the two applications involving medical image sequences cover some of the challenging properties investigated in section 2.2. Together, the presented MR and X-ray images provide a basis for both the evaluation of traditional image processing methods and the validation of the novel approach proposed in this work.

In the first section for each example application, solutions based on traditional image processing methods are investigated. The selected methods are low-level, intermediate-level, model-free as well as model-based. This selection covers different approaches and hence investigates the potentials and limitations of traditional methods. The conclusions drawn from these image processing solutions give a justification for a new approach and build a basis for comparison between the traditional and the novel approach.

5.2.1 Application to Carpal Bone MRI Sequences

5.2.1.1 Traditional Image Processing Approach

This section presents an image processing solution to the MRI application. The solution is based on classical methods to explore their potential and deficiencies. From a scientific point of view the development of a solution principle was comparatively easy, due to the restrictions that were imposed on the image formation process. These conditions allowed for an enhancement of the robustness and applicability of the solution and its implementation.

The SADT diagram in Figure 5.9 gives an overview of the overall system, which according to the structure presented in section 2.1.1 has been designed as a sequence of modules performing the following principle functions: image acquisition (A1), pre-processing (A2), and segmentation (A3) through feature extraction and matching. The measuring procedure has been divided into the measurement of the reference structure (A4) and also the objects of which the motion is to be determined (A5). A subsequent

module processes the measured parameter values to obtain motion graphs required to make a diagnosis.
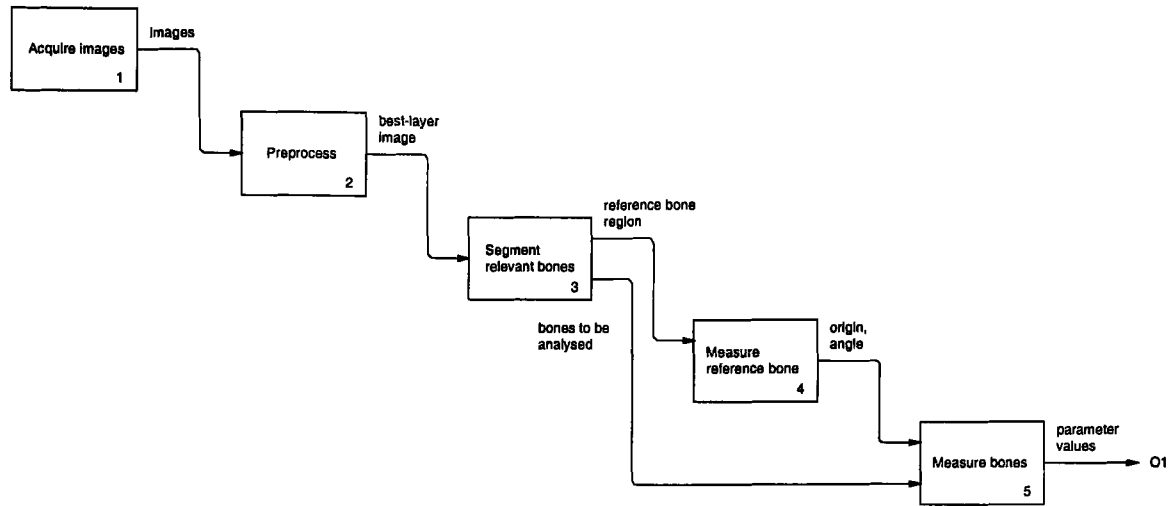
Figure 5.9: System structure of the traditional image processing solution to the analysis of carpal bones in MRI sequences.

Preprocessing

Due to the fixed parameter settings the MR image quality is good and almost constant, so that no noise reduction or contrast enhancement is necessary. The only preprocessing step that affects the pixel data is a grey-level normalisation that normalises the pixel values to a 10 bit range of $[0, 1023]$ since the grey-level range of the MR imaging device could not be fixed. 10 bits was empirically established as more than adequate for these images.

The two most important preprocessing steps involve the selection of a region of interest as well as the selection of a suitable MRI layer and are described in detail in the following subsections.

Automatic Selection of a Region of Interest

To allow for a histogram-based segmentation, a preselection of the area where the carpal bones are to be found is necessary (cf. Figure 5.10). This excludes metacarpals

and tissue which could be mistaken for carpal bones from the subsequent processing blocks. Furthermore, a major portion of the image is discarded considerably speeding up the pixel-based processing operations.

The region of interest (ROI) is found automatically through a simplified model of the hand based on anatomic *a priori* knowledge [156].

Figure 5.10: Example for the selection of a region of interest.

Layer selection

An approach based on the Fourier-Mellin transform [157] allows for the selection of the MR layer which is most suitable for the measurement. This involves comparing the set of MR input layers with a reference image, and a correlation measure is computed between a reference template selected by a medical expert and all slices of a hand position. The layer with the highest correlation factor is selected.[158]

Segmentation

An adaptive threshold is applied to the automatically selected region of interest. To obtain a higher precision the algorithm is applied in two stages to the ROI of the wrist and to smaller ROIs of the individual bones.

Firstly, a global threshold for all objects is found through an adaptive threshold algorithm based on the histogram of the ROI (cf. Fig 5.11a). The morphological operator erosion is applied subsequently to separate touching bone regions (cf. Fig 5.11b). The erosion which gives an unprecise object border is not critical, because the second segmentation step results in a more precise segmentation.

An identification (see next processing block) using an anatomic model removes all segmented regions that are not carpal bones, which allows the improvement of the segmentation of the carpals: for each bone an individual ROI is built within which a local adaptive threshold is determined.
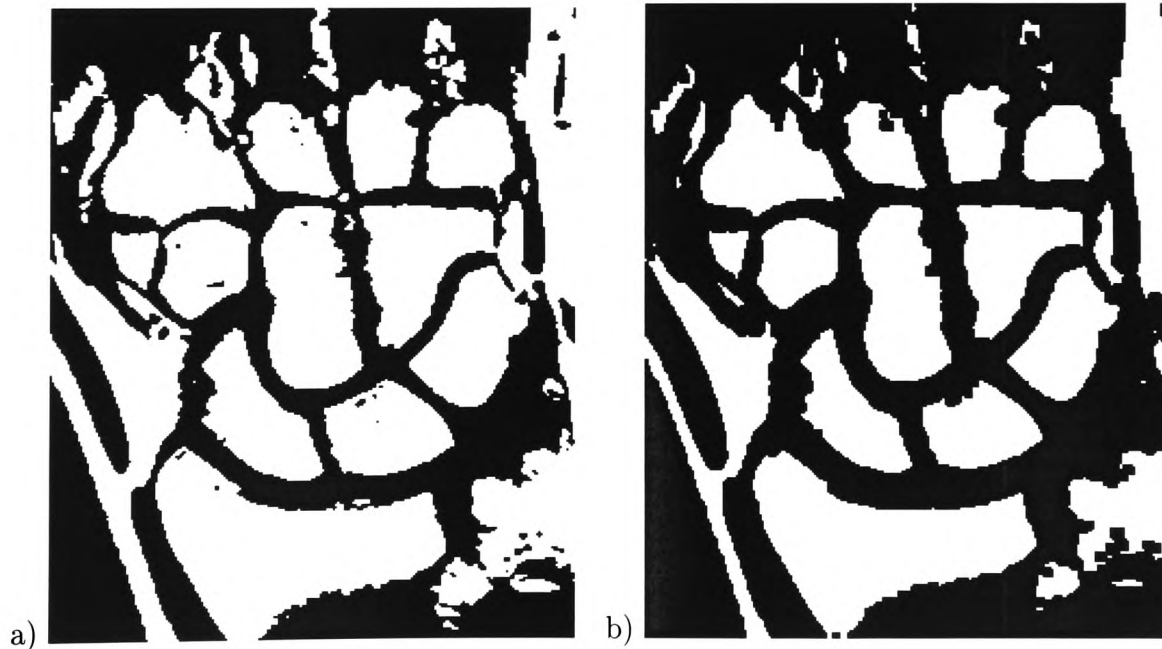
Figure 5.11: First-stage adaptive threshold a) before and b) after erosion to divide joined regions.

Identification

Apart from discarding unwanted regions from the subsequent local segmentation process the identification step is necessary to relate the measured parameters of all hand positions to the correct bones. As the reference co-ordinate system (cf. Figure 2.2) is based on landmarks of the Radius bone, this bone is output separately.

Constrained by their possible motion the relevant bones are identified through an analysis of the size and position of a set of candidate bones (cf. Figure 5.12).[159]
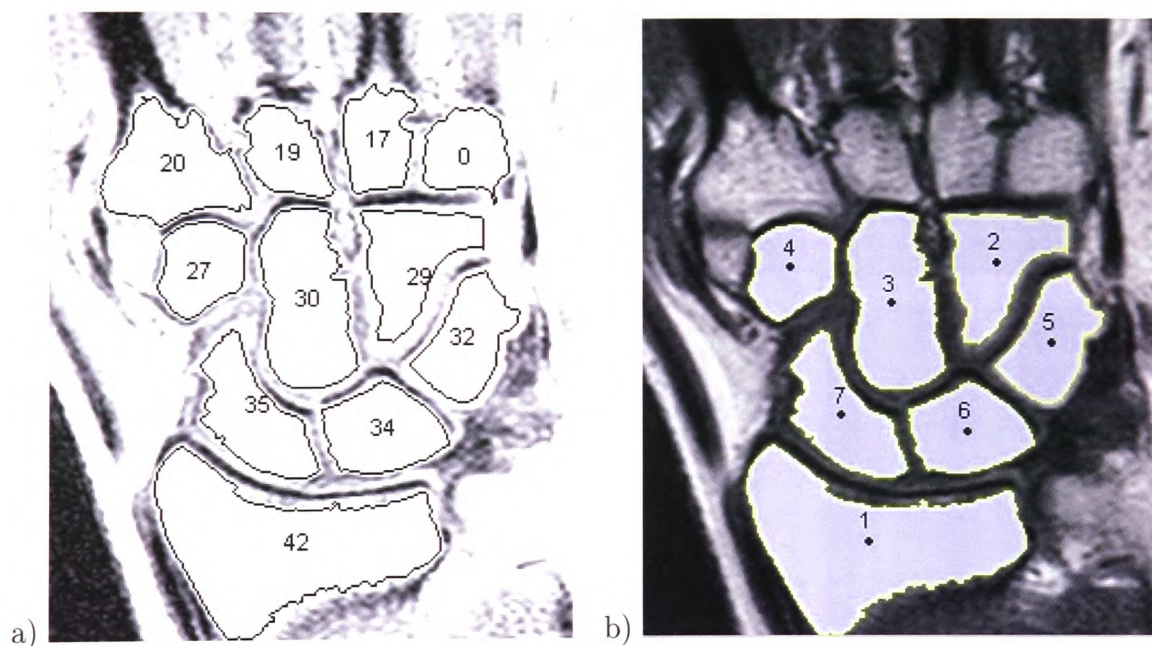
Figure 5.12: Identification stages. a) candidate regions, numbered randomly as found by the threshold stage. b) identified regions, numbered according to their anatomical meaning (cf. Figures 2.2 and 2.3).

Measurement

For the measurement of translation and rotation for each bone its major axis and centroid is determined. The measurement is performed relative to an anatomic co-ordinate system defined by the distal end of the Radius bone (cf. Figure 2.2).

Results

From an image processing point-of-view, successful segmentation is the most relevant factor to consider. The performance of the system is demonstrated by automatically measuring the motions of the bones of all together 158 wrist positions of 20 patients. A resulting number of 1106 bones were segmented. Figure 5.13 shows an example for the segmentation and measurement of all positions of a test person.

Figure 5.13: Segmentation and measurement results for all positions of one test person's hand. Images a) to i) correspond to hand positions of -34, -30, -20, -10, 0, 10, 20, 30, 40, and 48 degrees respectively.

Segmentation results were assessed by scientific and medical experts. Bone outlines were judged as being correct if the outlines included most of a bone's region and no extraneous tissue.

The segmentation was very successful for the most relevant bones *os scaphoideum*, *os lunatum* and the Radius (cf. Table 5.1). A good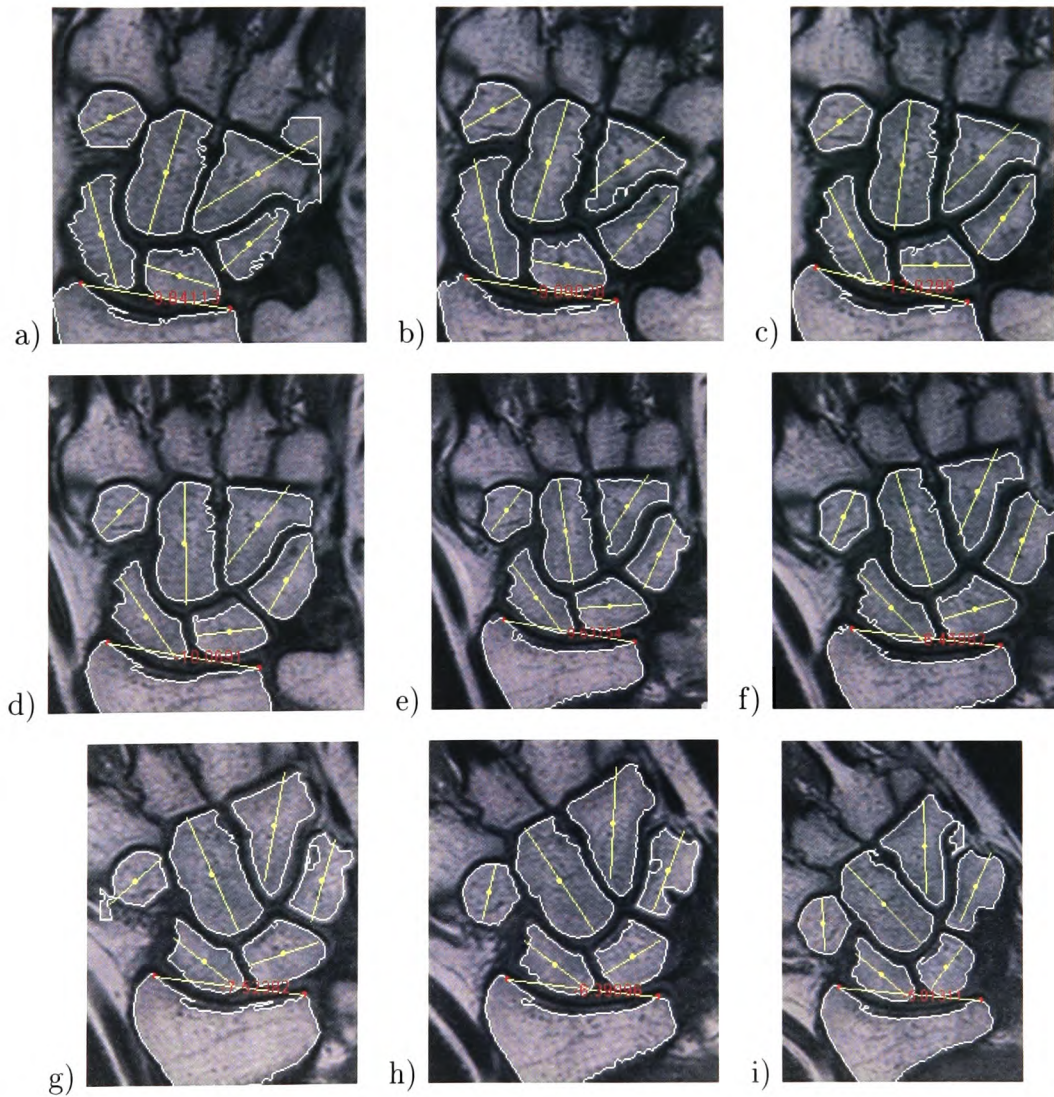 segmentation was also obtained for other carpal bones, allowing the system to be applied to the diagnosis of other carpal instabilities as well.

Bone	Correct segmentation [%]
os hamatum	77,8
os capitatum	94,9
os trapezoideum	90,5
os triquetrum	89,2
os lunatum	94,3
os scaphoideum	96,8
radius	97,5

Table 5.1: Correct segmentation rate.

Of particular importance for the overall goal of the application is to determine which of the measured parameters are significant to facilitate diagnosis.

Together with medical experts, motion graphs were developed, with the measurement results of usually 7 different positions of the wrist collected (cf. Figure 5.14). To ease comparison between normal and pathological bone motion, the rotation of each bone was normalised to the neutral position (an angle of 0 degrees) of the hand.

The bones' rotation was found to be particularly significant to aid in the diagnosis of carpal instabilities. In a healthy wrist there are two independent rows of bones, where the bones within each row are connected by ligaments: the distal row consisting of the *os hamatum*, *os capitatum* and *os trapezoideum*, and the proximal row including the *os triquetrum*, *os lunatum* and *os scaphoideum*.
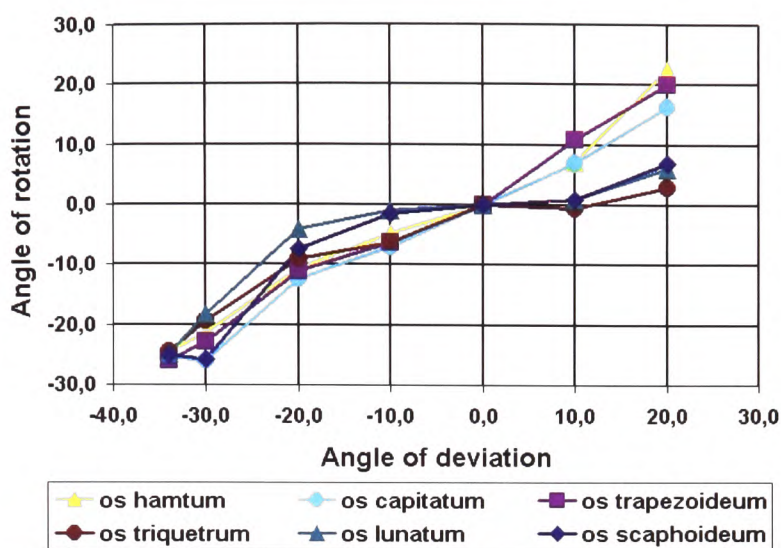
Figure 5.14: Motion graph — Collection of measurement results for several wrist positions (x-axis). This graph shows the rotation of the wrist bones (y-axis) for a healthy patient.
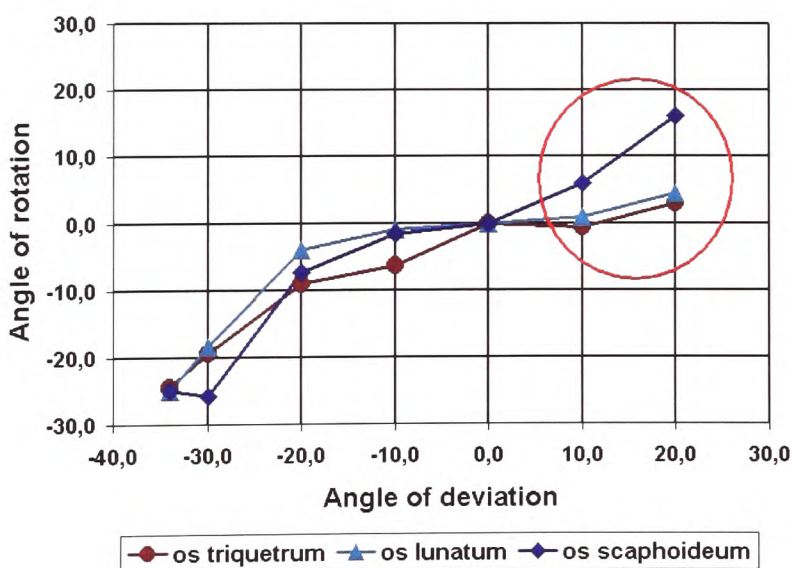
Figure 5.15: Motion graph, demonstrating the diagnosis of the scapho-lunate carpal instability.

Figure 5.14 reflects this normal state: It can be seen that the bones within each row move together over the full range of the wrist positions, while the rows themselves move independently for positive angles of the wrist.

The common lesion of the ligament between *os scaphoideum* and *os lunatum* (scapho-lunate instability) can be diagnosed with the aid of the motion graphs as demonstrated in Figure 5.15. For clarity the graph shows only the proximal row. Encircled is the range, where the motion of the *os scaphoideum* no longer follows the motion of the proximal row, indicating that the bone is no longer connected to the *os lunatum*.

The deficiencies of the traditional image segmentation approach are most evident with the *os hamatum*, which is the right bone of the distal row. It frequently touches neighboring bones or tissue. The bone also may have dents that are darker than the rest of the bone.

Figure 5.16: Segmentation results for the *os hamatum* with the traditional image processing approach.

A representative set of images was selected, where the traditional approach fails. Figure 5.16 shows nine images of five patients — three left and five right hands. For

comparison, the same set of images is used in the following sections, where an active contour and the proposed fuzzy snake are validated.

To detail the segmentation results outlined in Figure 5.16, the bone contour was divided into eight sections. The result of the assessment is given in Table 5.2.

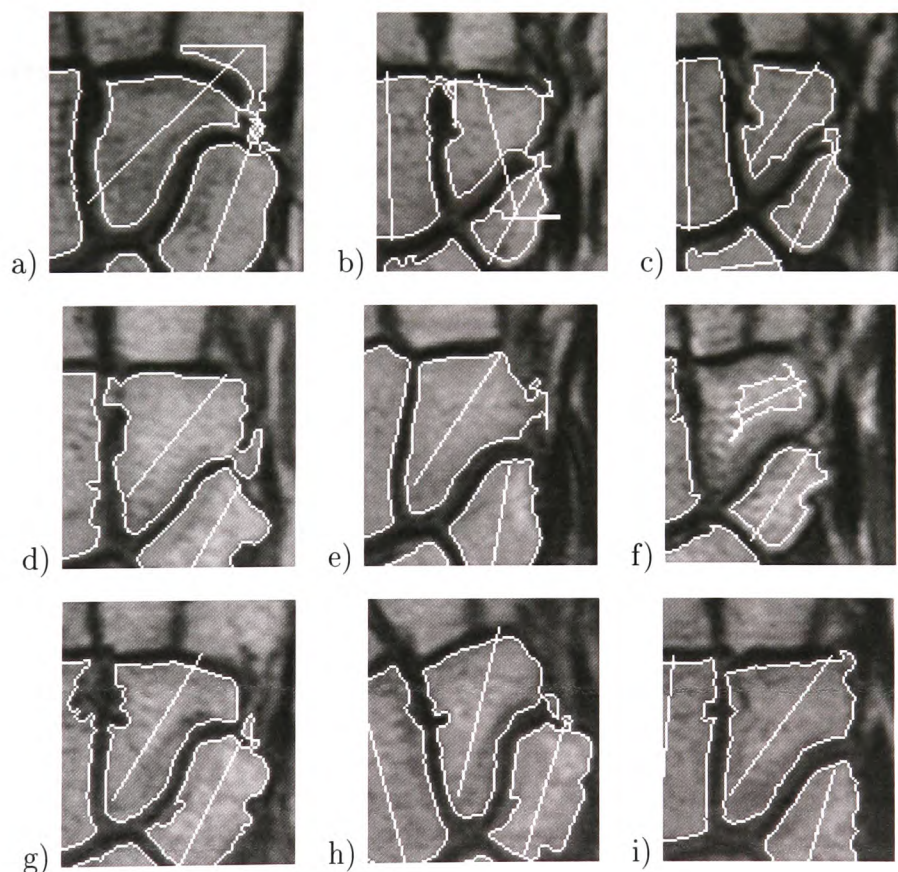| Correct contours	Correct sections
0/9	39/72 = 54%

Table 5.2: Correct segmentation rate for the *os hamatum* with the traditional image processing approach.

5.2.1.2 Traditional Active Contour Approach

This section summarizes an application based on an active contour as described in section 3.1. The particular energy functions applied in this solution were introduced in sections 4.3.3 and 4.3.3.2.

Traditional active contours use only global constraints, which improve the segmentation but this is not sufficient in certain cases. In Figure 5.17(e) for example, the global shape constraint is able to create a smooth contour resulting in the exclusion of most of the false tissue. The right-hand area of the result however, is still incorrect as the active contour is attracted to strong image features that could not be overridden by the moderate global shape constraint. A stronger influence of the constraint in this critical area is necessary.

As in the preceding section an assessment of the segmentation results (cf. Figure 5.17) was performed by both scientific and medical experts. The result of the assessment is given in Table 5.3.

Correct contours	Correct sections
1/9	47/72 = 65%

Table 5.3: Correct segmentation rate using a traditional active contour.
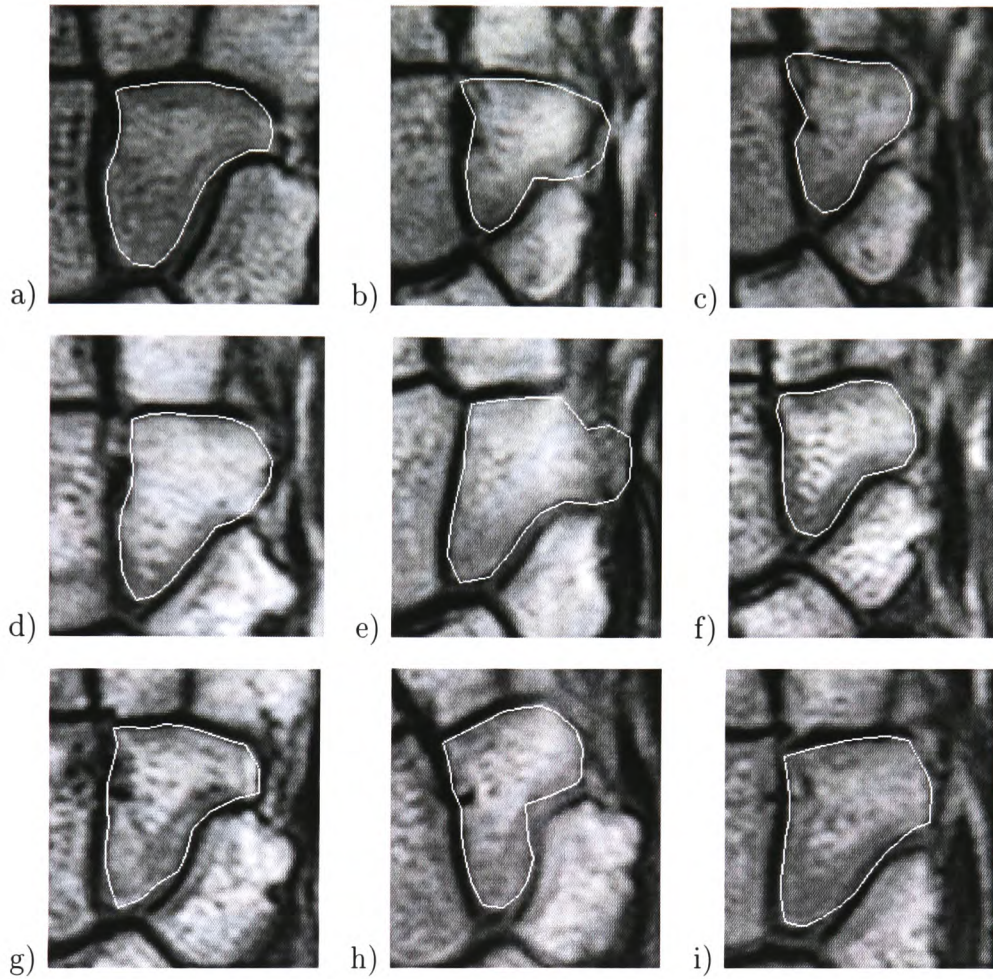
Figure 5.17: Segmentation results using a traditional active contour with global shape constraint.

5.2.1.3   Fuzzy Active Contour Approach

Fuzzy snakes allow for a more detailed modelling of the object, resulting in a correct segmentation. The results in Figure 5.18 were obtained with the description (using relative fuzzy lengths) $(0.2\,d)$ $(0.1\,b)$ $(0.05\,c)$ $(0.1\,e)$ $(0.15\,f)$ $(0.05\,c)$ $(0.05\,f)$ $(0.3\,a)$ beginning at the upper left vertex and following the contour clockwise. The characters are shortcut labels for the rule bases given in Table 5.5.

The fuzzy snake is able to handle a certain variability in the object's contour, which allows for the segmentation of a bone over a patient's MRI sequence as shown

in Figure 5.19. Furthermore the inter-individual variations are moderate, so the bones of other patients can be segmented successfully using the same contour description, or prototype [4]. The segmentation results were assessed in the same way as in the preceding sections. The result of the assessment is given in Table 5.4.

Correct contours	*Correct sections*
8/9	$71/72 = 98\%$

Table 5.4: Correct segmentation rate using fuzzy snakes.

Figure 5.18: Segmentation results using fuzzy snakes.

a:

IF *edge* rising medium

AND (*curvature* flat left

OR *curvature* flat right)

THEN *quality* very good

b:

IF *edge* rising medium

AND *curvature* medium right

THEN *quality* very good

c:

IF *edge* rising medium

AND *curvature* strong right

THEN *quality* very good

d:

IF *edge* rising strong

AND (*curvature* flat left

OR *curvature* flat right)

THEN *quality* very good

e:

IF *edge* rising strong

AND *curvature* medium left

THEN *quality* very good

f:

IF *edge* rising strong

AND *curvature* medium right

THEN *quality* very good

Table 5.5: Rule bases modelling the boundary segments of a wrist bone (*os hamatum*) in MRI sequences.

Figure 5.19: Fuzzy snake detection of two different positions with superimposed mean membership degrees for the *curvature* fuzzy sets. a)+c) highest and b)+d) second highest values.

5.2.1.4 Summary of Application to MRI Sequences

In this section three image processing and computer vision approaches were applied to medical MRI sequences introduced in section 2.3.1.

Histogram-based thresholding as well as morphological operators form the basis of a solution, which was presented to validate the potential of low-level methods. This approach produced satisfactory results detecting a number of medically relevant wrist bones. Where bones were touching soft tissue with a similar appearance however, the low-level approach failed.

Active contours employing simple shape constraints improved results but were also found to be incapable of producing a correct segmentation in many cases.

Only the improved representational power of the fuzzy active contour leads to a correct separation of bone and tissue.

Table 5.6 summarizes the performance of the validated approaches. The segmentation rates were obtained from an assessment by scientific and medical experts.

Approach	Correct contours	Correct sections
low-level	0/9	39/72 = 54%
active contour	1/9	47/72 = 65%
fuzzy snake	8/9	71/72 = 98%

Table 5.6: Correct segmentation rates of different approaches.

Different aspects of the traditional image processing solution were published in [2, 4, 5, 160-164]. The active contour approach was covered in [2, 4, 160, 163]. Additionally the application of the fuzzy active contour to the MR images was presented in [4, 160, 163].

5.2.2   Application to Vocal Tract X-ray Sequences

5.2.2.1   Traditional Image Processing Approach

In this section traditional image processing methods are investigated to analyse the human articulatory organs. This application is particularly well-suited to process different organs through different approaches from model-free low-level to intermediate model-based methods.

Preprocessing

The noise introduced by the scanning and digitisation process can be filtered out through low-level image processing. As the image formation has already produced blurry edges, the main concern with filtering is the preservation of edges. The literature recommends a median filter [21], the effects of which are shown in Figure 5.20. A median kernel size of 11x11 was empirically established as being able to reduce noise sufficiently (cf. Figure 5.20c). This family of nonlinear filters also have the added advantage of being able to remove small visible distortions due to scratches on the film.
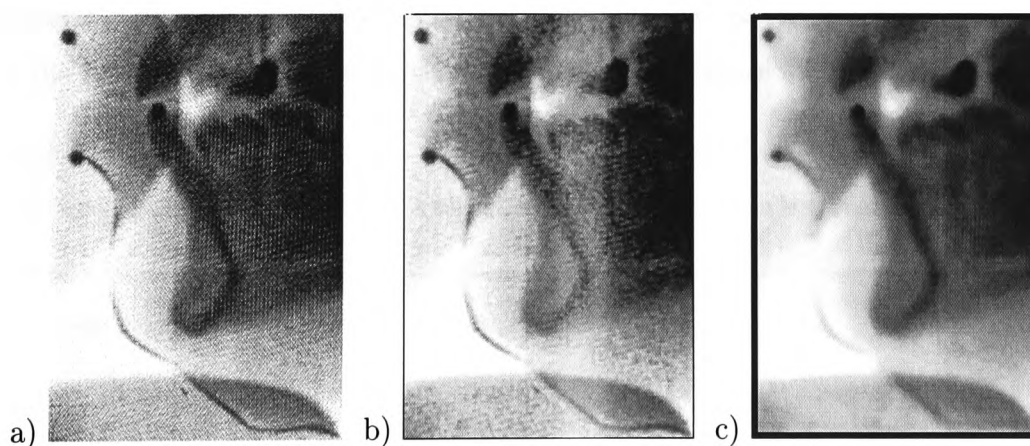

Figure 5.20: Application of a median filter with different kernel sizes to a region of an X-ray image: a) original, b) median filtered (3x3 kernel), c) median filtered (11x11 kernel)

Segmentation of rigid objects

A low-level image segmentation method based on intensity statistics demonstrates the principle of such an approach and the results which can be obtained with the presented X-ray sequences. As an example for a rigid object with a relatively good contrast the front teeth (cf. Figure 2.8) were selected.

Generally a local edge detecting operator based on gradient approximation attenuates not only edges of the object under investigation but also any other structures in the image (cf. Figure 5.22a). Hence the gradient value $\nabla I(x, y)$ of a pixel represents its probability of being a contour pixel *(edge probability)*, but it does not state which object the pixel belongs to. The basic concept of the approach is to multiply the gradient value with a transformed intensity value which represents the probability of a pixel to belong to the object contour under investigation *(object probability)*. The result of this operation is a pixel's probability of being a contour pixel of the desired object.[17, 165]

With blurred edges there is no sharp transition between object and background. Hence object contours cover a relatively wide area of pixels and a range of intensity values rather than a single pixel with only one intensity. Consequently it is possible to calculate a histogram for contour pixels of a reference object to obtain information about the intensity characteristics of that contour (cf. Figure 5.21).

While the edge probability is a monotonic function of the gradient value this does not hold for object probability and contour histogram. Thus an intensity transformation is necessary to obtain a similar relationship. The enclosing hull of the discrete histogram $p_C(i)$ can be approximated by the Gaussian

$$f_G(i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(i-\mu)^2}{2\sigma^2}} \tag{5.1}$$

where the intensity values $i \in [0, Q - 1]$.
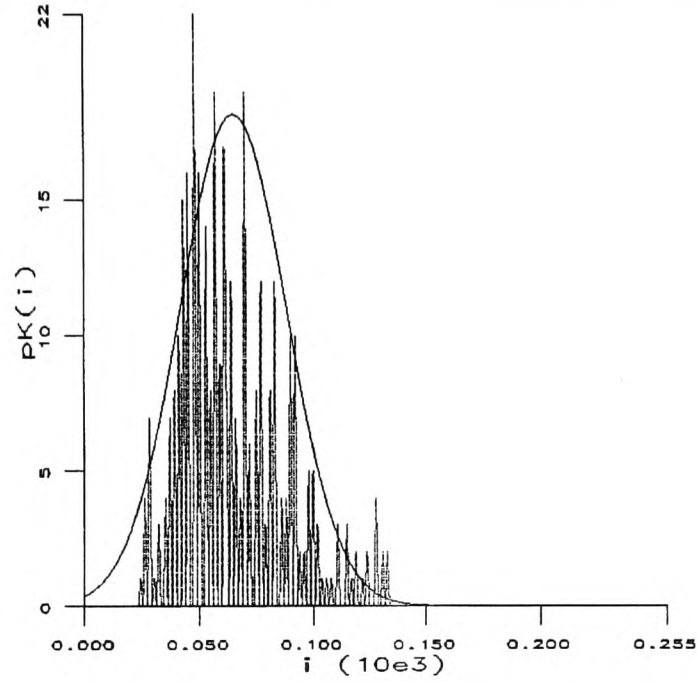
Figure 5.21: Intensity distribution $p_C(i)$ of a reference contour. Superimposed is the Gaussian approximation, characterised by μ and σ^2 which were obtained from $p_C(i)$.

The parameters variance σ^2 and mean μ can be obtained from the histogram $p_C(i)$ through

$$\mu \;=\; \sum_{i=0}^{Q-1} i\, p_C(i) \tag{5.2}$$

$$\sigma^2 \;=\; \sum_{i=0}^{Q-1} (i - \mu)^2\, p_C(i) \tag{5.3}$$

After sampling and normalising the Gaussian $f_G(i)$ to $f_{Gn}(i)$ so that $f_G(\mu) = 1$, the intensities of the input image I are transformed (cf. Figure 5.22b):

$$\forall (x,y) \,:\, I_T(x,y) = f_{Gn}(I(x,y)) \tag{5.4}$$

Multiplying the gradient image ∇I with the intensity transformed I_T results in a feature matrix (cf. Figure 5.22c) from which the contour pixels can be obtained using a threshold (cf. Figure 5.22d).
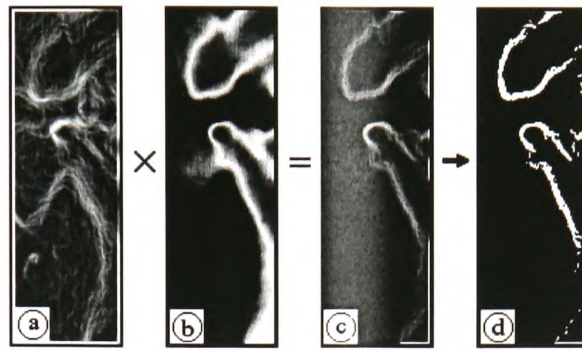
Figure 5.22: Tooth segmentation: a) gradient ∇I, b) intensity transformed input image I_T, c) resulting feature matrix, d) contour after thresholding.

The above method is an example for a low-level segmentation approach based on local grey-levels as well as on a statistical model of the object contour's intensity values. Example results are shown in Figure 5.23. They demonstrate some of the deficiencies of the approach, which are:

- As the grey-levels of reference contour points are used to classify the pixels, the method is not robust against intensity changes that differ from the reference.

- Unwanted objects whose contour intensity range is similar to the reference are also segmented. Examples are leaden markers, other teeth, fillings or dirt on the celluloid.

- The resulting contour is not always continuous and closed.
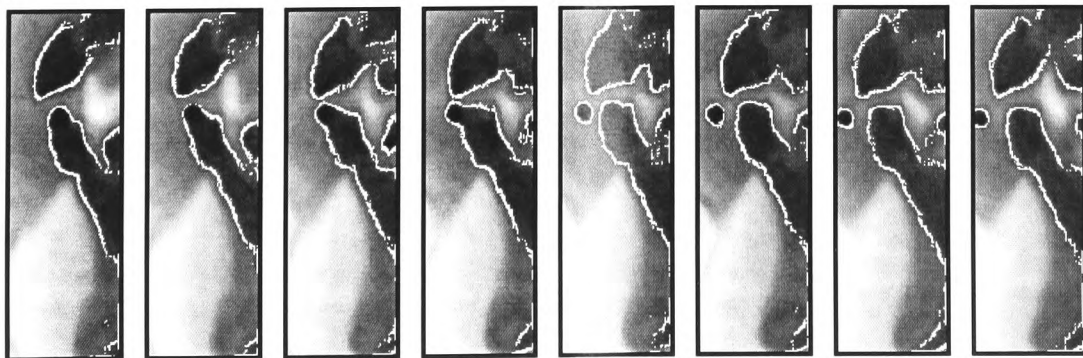
Figure 5.23: Results for tooth segmentation.

A different low-level image processing method is template matching. Here a template image containing a reference object image (cf. Figure 5.24b) is successively correlated with all areas of the larger source image (cf. Figure 5.24a). The maximum correlation value indicates the best matching location of the template within the source.
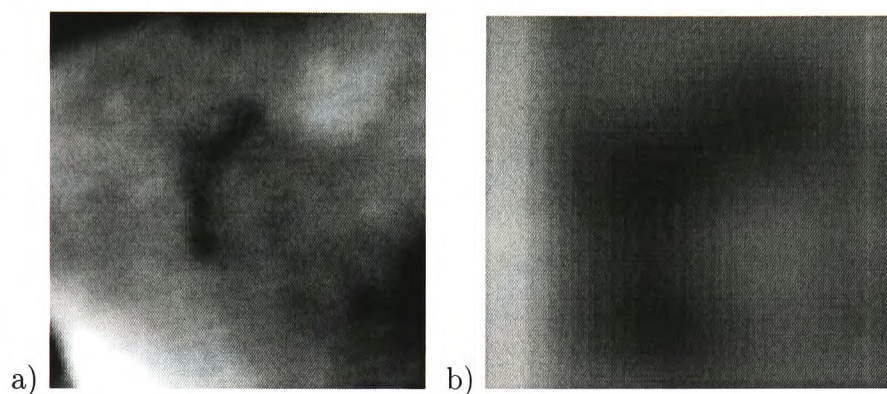
Figure 5.24: a) Hyoid bone ROI after preprocessing, b) template (enlarged).

An advantage over the previously investigated method is that a normalised correlation measure may be used. This is invariant against moderate changes in intensities. Furthermore the detection of false objects is less likely, as the template contains not only intensities but also their position in the image.

The hyoid bone is an object where this approach can be applied (for more details see [166]). It is rigid and mostly free from superimposing objects. Figure 5.25 shows some examples for a successful location of the hyoid bone.
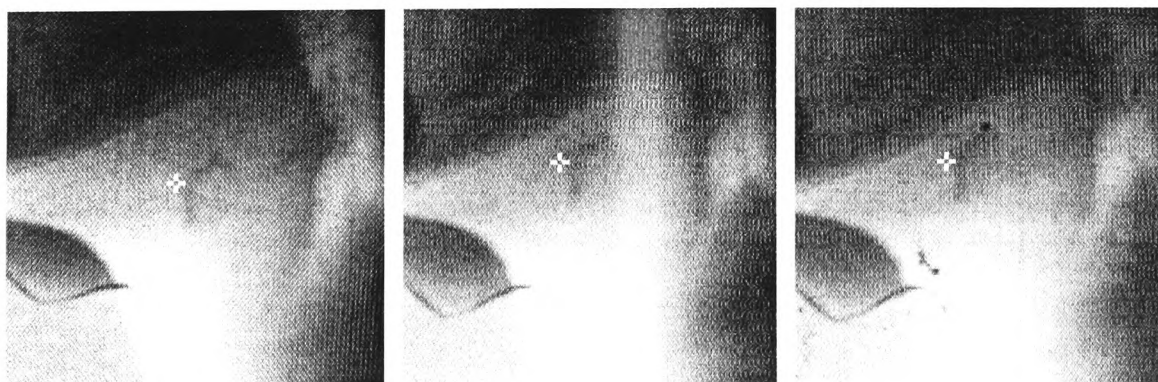
Figure 5.25: Successful detection of the hyoid bone.

Although the template matching approach was highly successful for the location of the hyoid bone, it cannot be applied to the other articulators for the following reasons:

- The applied correlation is not invariant against scaling and rotation. Two invariant template matching approaches [157, 167] have been investigated in [158, 168] and proved to be promising. An in-depth exploration however, is beyond the scope of this thesis.

- The image correlation approach directly outputs the co-ordinate of an object point rather than a contour. It is therefore not suitable for the determination of the parameters listed in Table 2.5.

- Template matching is not robust against occlusion and deformation.

Segmentation of deformable objects

An intermediate-level approach is investigated next to detect deformable objects such as the lips and the tongue. The method uses *a priori* knowledge about the object to be detected on a higher level than the intensity-based approaches presented above. It allows for the exploitation of simple geometrical constraints and guarantees an uninterrupted one-pixel wide contour. The approach developed by [51] and implemented in [165] involves a constrained search algorithm which traces a contour pixel by pixel within a region of interest (ROI) close to the actual object contour.

Initially the ROI is determined by manually marking the desired object contour in the first image of a sequence (cf. Figure 5.26a). This contour is referred to as the initial contour hypothesis, as it represents an initial estimate for the actual position of the contour in the subsequent image. The width of the ROI is a system parameter and depends on the maximal distance the object might travel from one frame to another. When the object has been segmented, the resulting contour serves as initial estimate for the next frame, hence tracking a moving or deforming object.

Such a region of interest leads to a significant reduction of the search space and hence improves the computational complexity of the segmentation. More importantly it introduces implicit knowledge to the system about where the object is to be found. Consequently distortions such as other objects do not impair the segmentation as long as they are outside the ROI. Furthermore knowledge about the approximate shape is also introduced.

To simplify the segmentation algorithm, the ROI is transformed into a straight matrix (cf. Figure 5.26b). Assuming that the initial contour hypothesis is approximately parallel to the object contour and that parts of the ROI do not overlap, the straight matrix contains exactly one contour point per row. The contour tracing algorithm now processes each row to find the optimal horizontal position of each contour point. After all rows have been processed the detected contour points are transformed back into the original image to obtain the final contour.
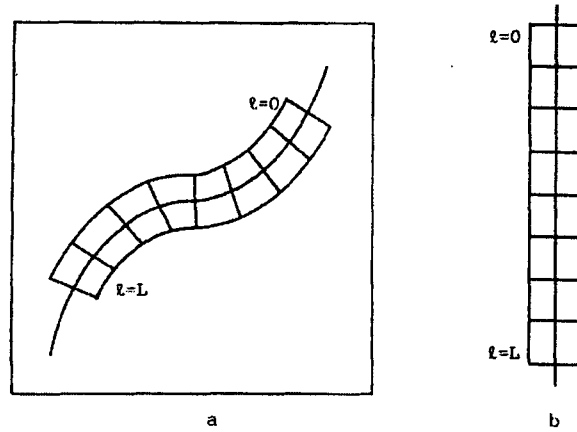
Figure 5.26: a) Initial contour hypothesis with ROI in the source image, b) straight matrix after geometrical transformation of the ROI. From [51].

To find the optimal contour point in each row a dynamic programming algorithm is applied. Dynamic programming in general is described in [125]. Other contour segmentation methods based on this optimisation principle are presented in [16, 20, 126, 127]. As dynamic programming is also used in the novel approach presented in this thesis it is investigated thoroughly in section 3.1.3.2.

Of more relevance to the current section is the image feature that is used to segment the contour. With the segmentation of rigid objects the use of the image gradient was presented as a method to detect edges. With the transformed ROI the gradient can be used more effectively. As in each row the detection of the optimal contour point is reduced to a one-dimensional problem, the (horizontal) gradient can be approximated by a one-dimensional convolution[18, 21], the coefficient matrix of which is

$$H_h = [-1 \quad -2 \quad 0 \quad 2 \quad 1] \tag{5.5}$$

Edges parallel to the initial contour hypothesis will be attenuated, while edges perpendicular to it will be suppressed. This directional sensitivity represents an edge detection based on *a priori* knowledge. It is an advantage over the two-dimensional edge detection which applies a directional insensitive operator to the original image.

As the horizontal edge orientation is parametrised by the sign of the coefficient matrix, the number of possible contour points can be reduced further. Applying H_h denotes a transition from dark to bright intensities (when viewed from left to right), while $-H_h = [1 \quad 2 \quad 0 \quad -2 \quad -1]$ will detect bright-to-dark transitions. This allows for an implicit inclusion of knowledge about the inside or outside of objects which do not change their brightness relative to the background.

Figure 5.27 shows an example sequence where the lower lip was successfully segmented using the contour tracing approach. Due to the local and geometric restrictions the segmentation is robust against distortions that occur within the shown image region but outside the ROI around the contour. The directional gradient ensures that the contour is not attracted by for example the upper lip. Finally the algorithm follows small deformations of the lip.

When applied to the tongue the limits of the approach become evident. In Figure 5.28 the contour roughly follows the tongue in most parts. Particular difficulties arise in the front part, where the contrast and hence the gradient is very low. Consequently the contour locally deviates from the optimum. This deviation becomes larger over the course of a sequence and eventually affects other parts of the contour, as the initial contour hypothesis becomes more and more inaccurate.
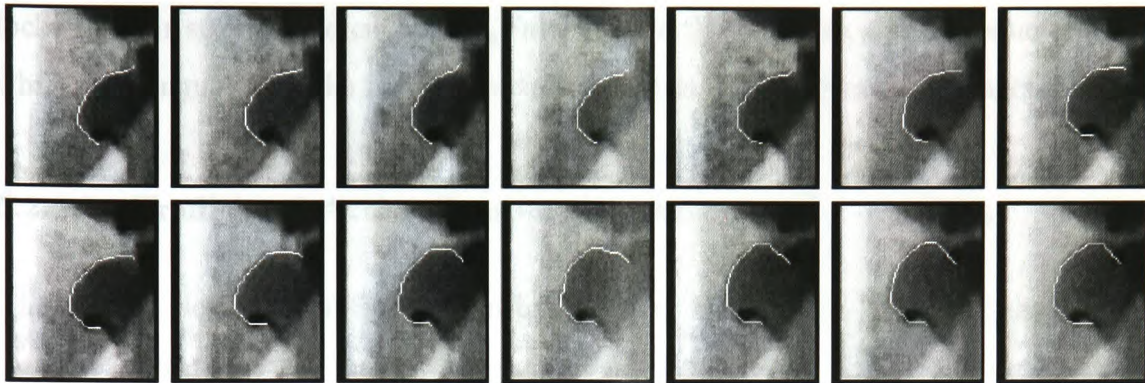
Figure 5.27: Results of a dynamic-programming based contour tracing system using a minimum of *a priori* knowledge about the object. The images show the region of the lower lip in successive frames (from left to right and top to bottom), with the identified contour superimposed.

Figure 5.28: Insufficient tongue segmentation by local contour tracing.

The most important deficiency of the algorithm is its lack of robustness against occlusion and varying texture[1]. When the tongue moves near teeth or fillings, the contour

---

[1] A deeper investigation of texture-based segmentation methods is beyond the scope of this thesis. General information is provided by image processing text books [8, 13–21]. In addition [169] presents a texture-based approach to find ill-defined objects.

locks to their strong gradient values. Furthermore the resulting contour is not smooth while the tongue is, as the flowing contrast agent is not distributed homogeneously.

5.2.2.2 Traditional Active Contour Approach

This section describes an application based on an active contour as described in section 3.1. The particular energy functions applied in this solution were introduced in sections 4.3.3 and 4.3.3.2.

Contour Identification and Measurement System

Figure 5.29 shows the overall structure of the proposed modularised contour identification and tracking system.
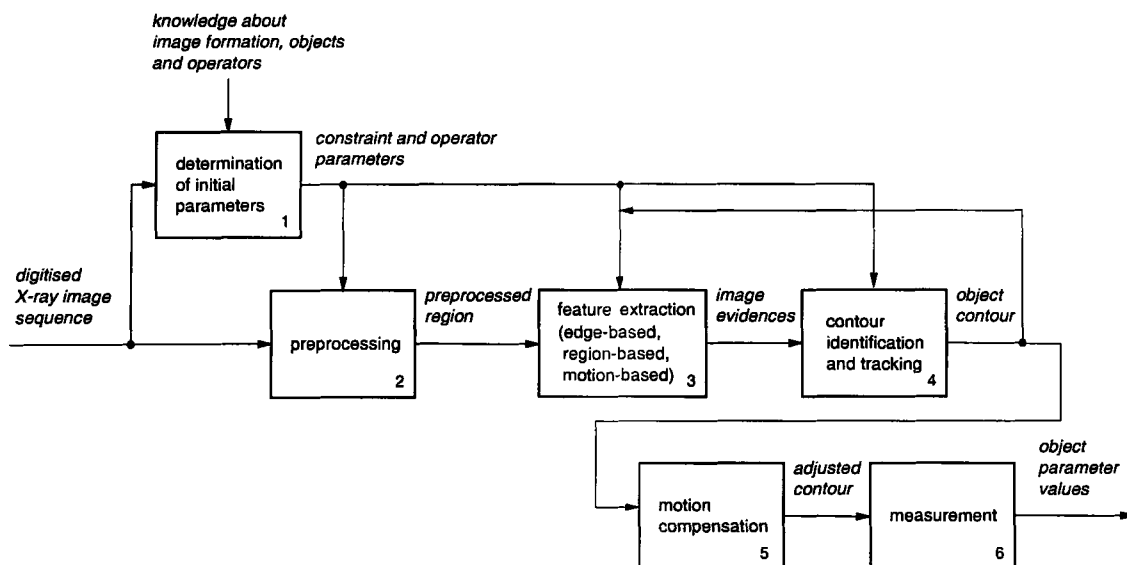
Figure 5.29: System structure.

The system's characteristics are:

- A sequence is processed frame by frame. Currently all objects are processed independently.

- Except for the first image of the sequence, where the relevant contours have to be marked by a human expert, the system is fully automatic.

- Multiple image evidences from different image operators are integrated.

- The detected contour is fed back into the system to improve detection in the next frame.

Feature Extraction

The task of this module is to extract image features e_{image} that form the basis for the subsequent contour identification module. Unlike other approaches the proposed module does not segment tokens [10], such as geometric objects like lines or polygons, and either group or delete them afterwards based on their individual features. Instead, the features remain pixel-related in order to be able to identify a precise contour, as outlined in section 4.3.3.1.

Figure 5.30 shows the tongue's search space **R** as defined in Figure 3.1. The feature extraction is applied to this search space, or region of interest, according to equations (4.4) to (4.7) (cf. Figure 5.31).
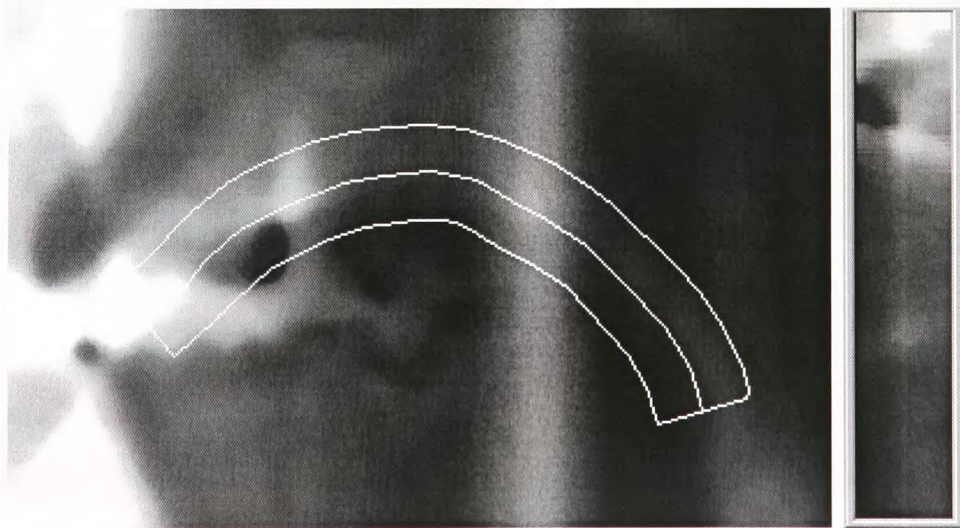
Figure 5.30: The search space **R** is derived from an initial contour hypothesis (centre line).
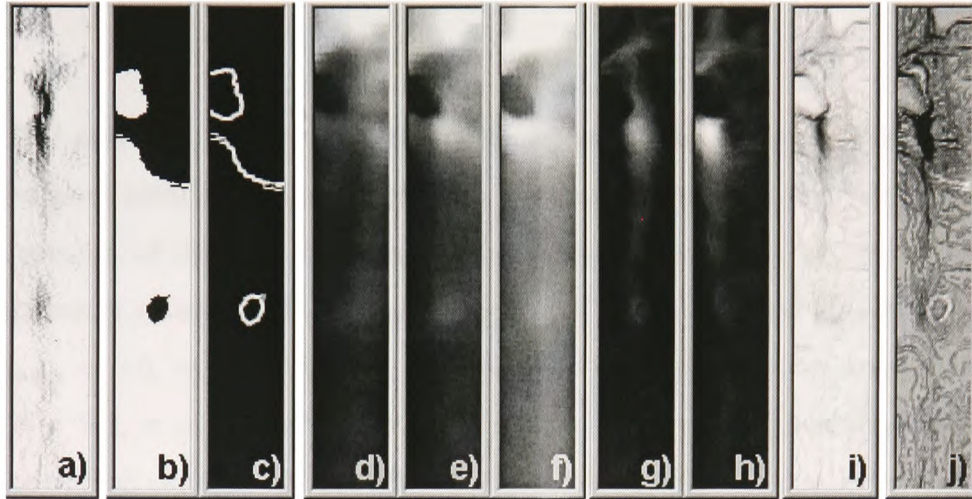
Figure 5.31: Sub-components of e_{image}. a) e_{edge}, b) Θ ($\theta = 80$), c) e_{region}, d) $\mathbf{R}_{t-d_1}$ ($d_1 = 5$), e) $\mathbf{R}_t$, f) $\mathbf{R}_{t+d_2}$ ($d_2 = 5$), g) $|\mathbf{R}_{t-d_1} - \mathbf{R}_t|$, h) $|\mathbf{R}_t - \mathbf{R}_{t+d_2}|$, i) e_{motion}, j) e_{image} ($w_{edge} = 0.1, w_{region} = 0.1, w_{motion} = 0.8$)

Constraints

A shape constraint was presented in [1] and described in section 4.3.3.2 which improves the detection of the tongue considerably (cf. Figure 5.32).
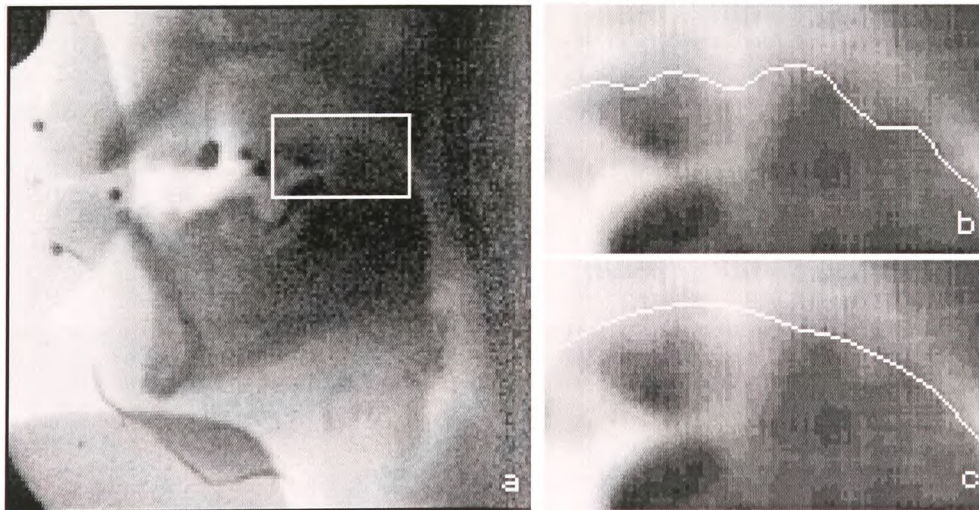
Figure 5.32: Examples of contour detection without constraint (b), and with the proposed shape constraint (c).

Results

Figure 5.33 and 5.34 demonstrate the performance of the proposed system, showing some results of identifying the tongue in a number of frames of two X-ray sequences. Both sequences of different test persons have been successfully processed using the same parameter settings. The weighting factors of the normalised energy components were $w_{image} = 0.6$, $w_{int} = 0.15$ and $w_{con} = 0.25$. The constraint parameters were set to $\bar{\gamma} = \pi - 0.2$, $\sigma = 0.2$, $\gamma^{min} = \pi - 0.4$ and $\gamma^{max} = \pi$. These settings were easily found since small changes did not affect the results. The parameters of the geometrical constraint proved to be particularly robust.

The tongue is partially occluded by teeth, cheek bones and fillings. Due to salivation the distribution of the contrast agent which was applied orally, is not constant. Although these influences lead to varying features of the boundary, the tongue is detected correctly.
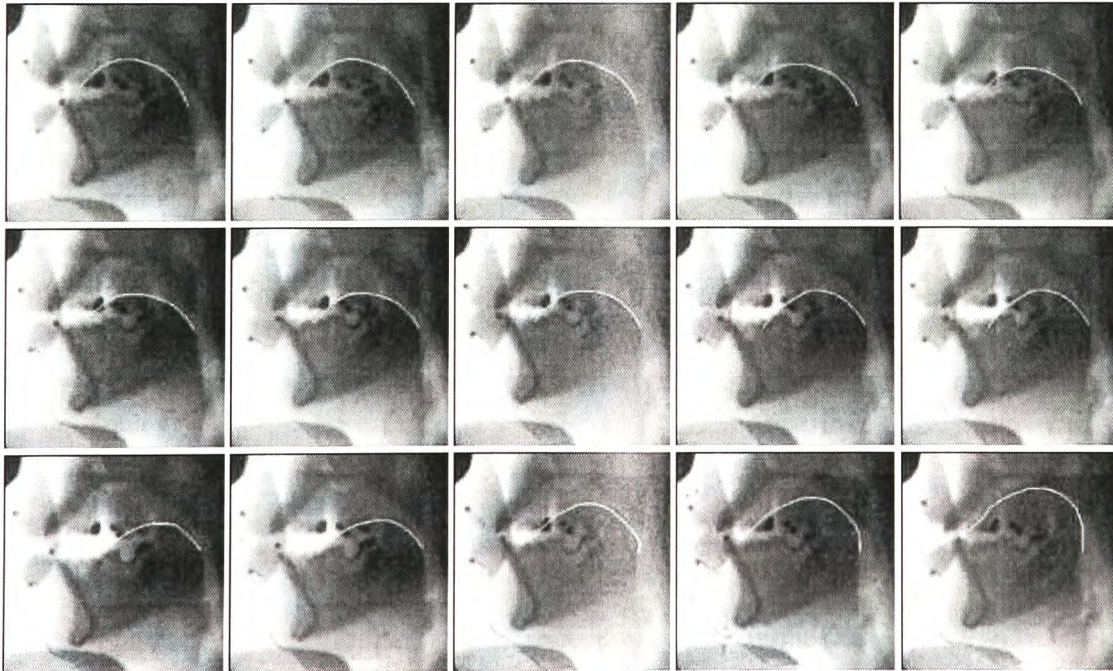
Figure 5.33: Results of the proposed system identifying the tongue. Every second frame of a sequence is shown.

Even when the back tongue touches the velum so there is little distinction to be made between the two articulators, the active contour correctly aligns with the boundary of the tongue. Figure 5.34 demonstrates the same robustness when the middle tongue touches the soft palate. This result is of particular importance to the analysis of speech production, as the location and diameter of constrictions constitute important parameters to characterise a particular sound.

Figure 5.34: Results of the tongue tracking system on a different probationer. Every third frame of a sequence is shown.

The final image in Figure 5.33 shows a concave section in the front tongue, with the boundary correctly aligned with salient image features. This demonstrates the ability of the system to locally violate the geometrical constraint under the presence of strong image evidences.

With the global shape constraint however, it was not possible to segment the tip of the tongue, which has a significantly higher curvature than the rest of the tongue.

Further limitations of the global constraint become evident when the front tongue bends to a distinct concave shape. The active contour then aligns to false image evidences in order to satisfy the shape constraint.

To demonstrate this effect and to compare this approach with the results of the proposed fuzzy snake, a characteristic set of images was selected. Figures 5.35 a)-d) show four images, where the tongue's shape was adequately modelled with the active contour's global constraint. In Figure 5.35 e) the tip of the tongue starts to rise, violating the shape constraint in the front section. In the subsequent images in Figures 5.35 f)-i) the tip further bends towards the front teeth and alveolus.
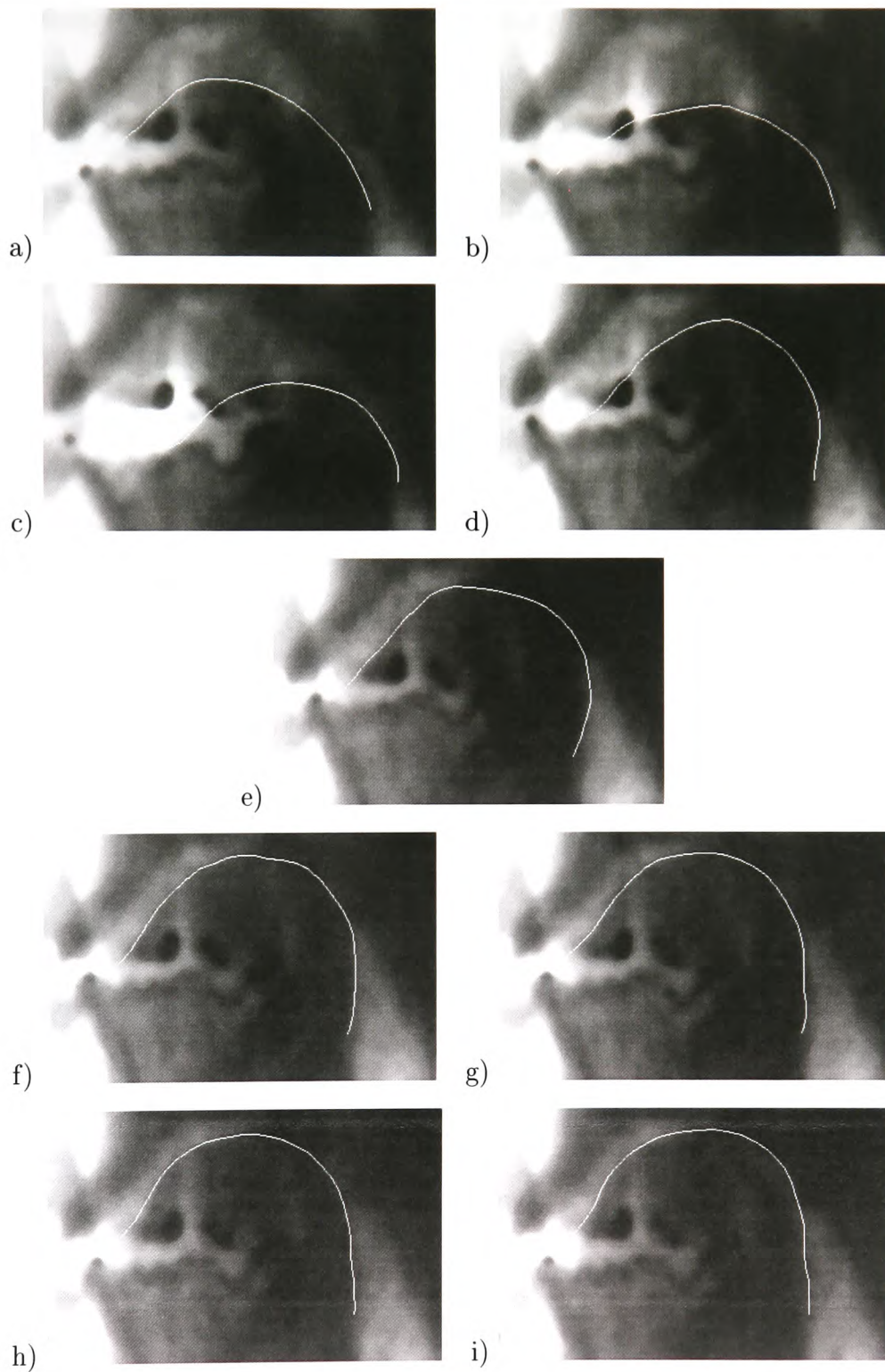
Figure 5.35: Segmentation results using a traditional active contour with global shape constraint.

An assessment of these segmentation results was performed by scientific experts. To detail the assessment, the tongue was divided into the four sections tip, front, middle and back. The result of the assessment of Figure 5.35 is given in Table 5.7.

Correct contours	Correct sections
0/9	18/36 = 50%

Table 5.7: Correct segmentation rate using a traditional active contour.

5.2.2.3 Fuzzy Active Contour Approach

Satisfactory results for many positions of the tongue were presented (cf. Figures 5.33 and 5.34), emphasising the need for combined image features as well as a shape constraint as outlined in sections 4.3.3.1 and 4.3.3.2.

To detect different shapes of the tongue it was necessary to allow for a relatively high tolerance in the shape constraint, but this however reduced the robustness of the active contour against distortions. Furthermore highly concave sections of the front tongue as well as the tip of the tongue were not detected as the locally different curvature could not be modelled through the global shape constraint.

These problems were overcome using the more detailed contour description afforded by the fuzzy snake. The description in Table 5.8 together with the rule bases in Table 5.10 allowed for a modelling of the segments of the tongue for characteristic positions.

very short tip, *short* front, *medium* bent middle, *medium* bent back

Table 5.8: Contour description of the tongue.

Locally different curvatures were accounted for, increasing the robustness of the segmentation as well as enabling the detection of the tip of the tongue (cf. Figure 5.36). Structures that were able to distort the snake were explicitly modelled, increasing the correctness of the segmentation. The second rule in the rule base tip for instance,

accounts for the lower jaw while the IF-*region* rule in the rule base **front** reduces the influence of a tooth filling located in the upper jaw.

The segmentation results shown in Figure 5.37 were assessed in the same way as for the active contour approach. The result of the assessment is given in Table 5.9.

Correct contours	*Correct sections*
8/9	35/36 = 97%

Table 5.9: Correct segmentation rate using using fuzzy snakes.
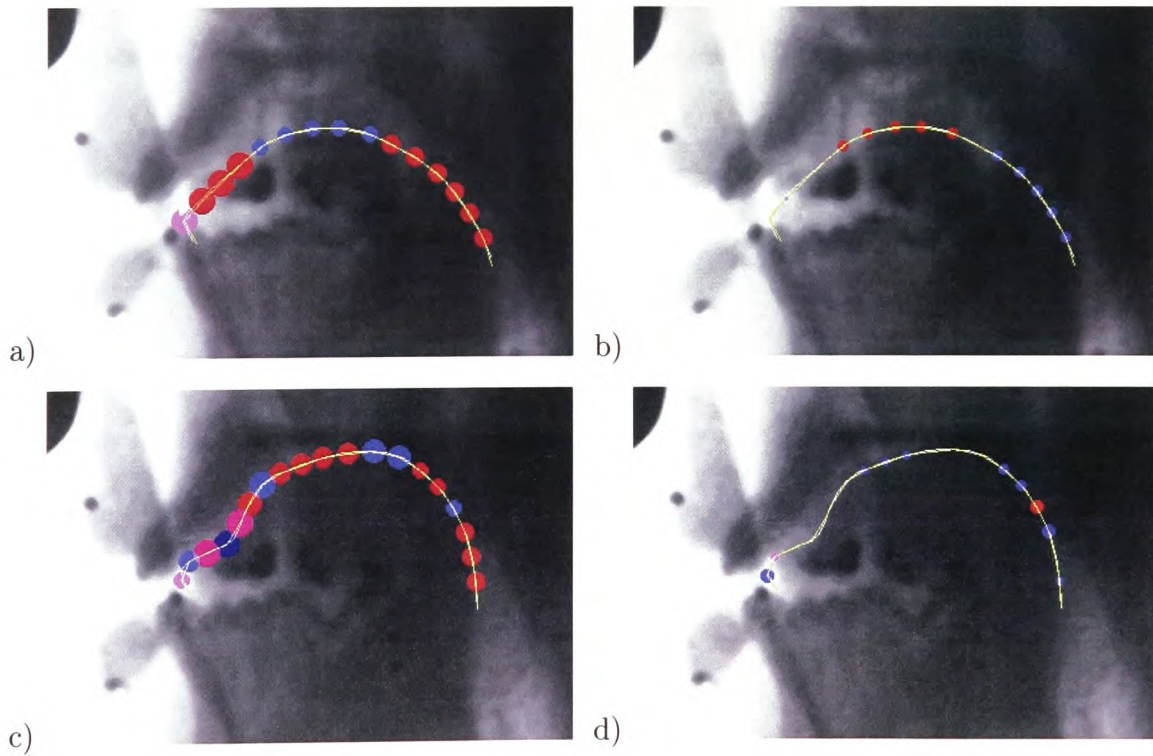
Figure 5.36: Fuzzy snake detection of two characteristic positions of the tongue with superimposed mean membership degrees for the *curvature* fuzzy sets. a)+c) highest and b)+d) second highest values.
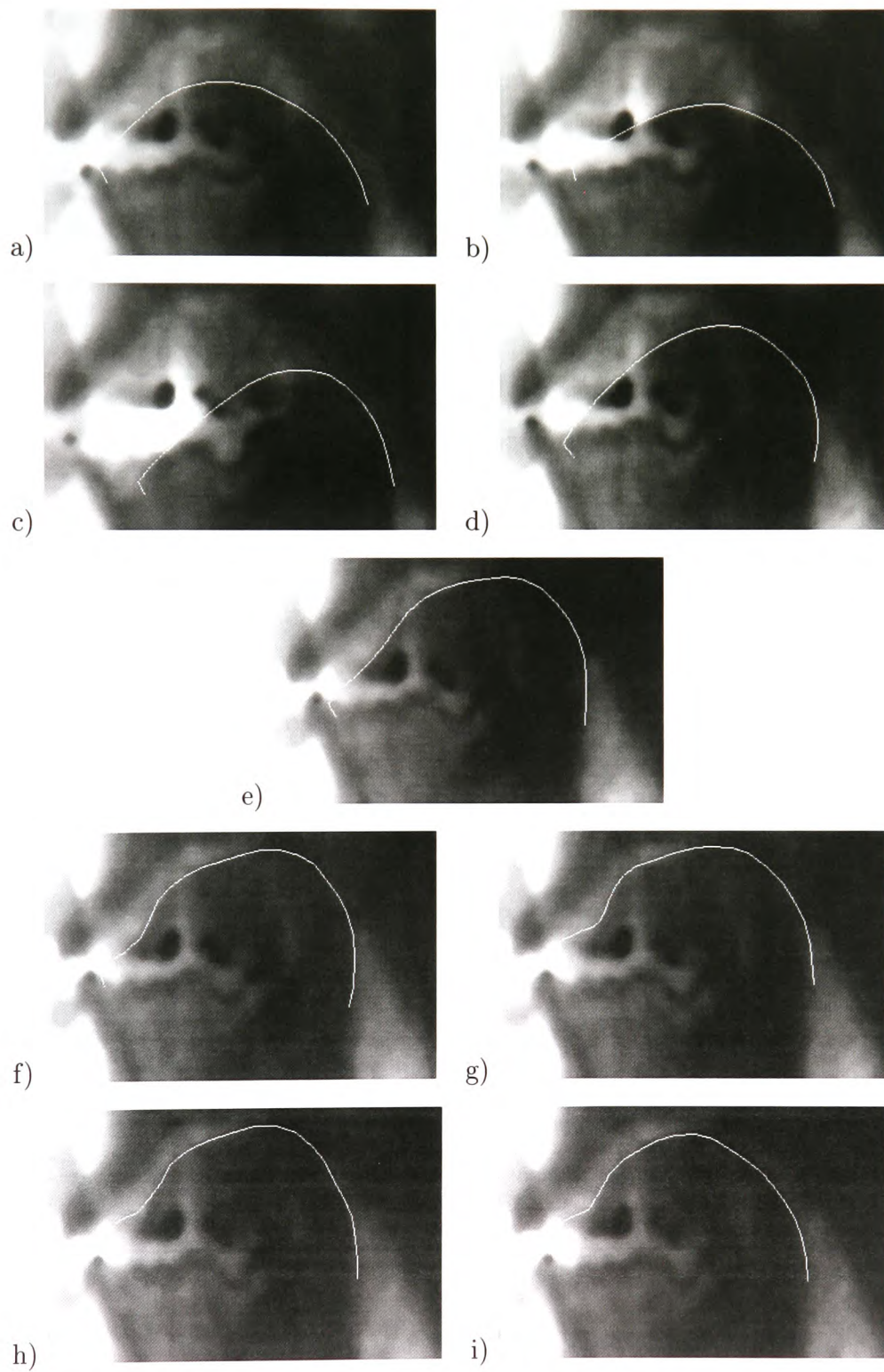
Figure 5.37: Segmentation results using fuzzy snakes.

tip:

IF *edge* falling very weak

AND (*motion* negative medium

OR *motion* positive medium)

AND *curvature* strong right

THEN *quality* very good

IF *edge* falling medium

OR *edge* falling strong

THEN *quality* bad

bent middle:

IF *edge* falling weak

AND (*motion* negative strong

OR *motion* positive strong)

AND (*curvature* flat right

OR *curvature* medium right)

THEN *quality* very good

bent back:

IF *edge* falling medium

AND (*motion* negative medium

OR *motion* positive medium)

AND (*curvature* flat right

OR *curvature* medium right)

THEN *quality* very good

front:

IF (*edge* falling medium

OR *edge* falling weak)

AND (*motion* negative strong

OR *motion* positive strong)

AND *curvature* flat right

THEN *quality* very good

IF (*edge* falling medium

OR *edge* falling weak)

AND (*motion* negative strong

OR *motion* positive strong

OR *motion* negative medium

OR *motion* positive medium)

AND *curvature* medium right

THEN *quality* very good

IF (*edge* falling medium

OR *edge* falling weak)

AND (*motion* negative strong

OR *motion* positive strong

OR *motion* negative medium

OR *motion* positive medium)

AND (*curvature* flat left

OR *curvature* medium left)

THEN *quality* very good

IF *region* negative strong

OR *region* negative very strong

THEN *quality* bad

Table 5.10: Rule bases for modelling the boundary segments of the tongue.

5.2.2.4 Summary of Application to X-ray Sequences

This section has initially presented traditional image processing solutions to validate the potential of low-level and intermediate-level methods applied to medical X-ray sequences introduced in section 2.3.2.

It was verified that low-level image processing based on intensity statistics and local edge detection can detect rigid objects with a high local contrast. Where the contrast is low and edges blurry, more advanced low-level techniques such as template matching can be successful.

Both model-free low-level techniques fail for partially occluded and deformable objects. It then was shown that such variations require *a priori* knowledge, which in turn requires an at least intermediate-level approach. The investigated contour tracing approach models the objects to be detected through a region-of-interest and a preferred edge direction.

This knowledge on shape and appearance is however merely implicitly represented by the approach, making its application and extension difficult. Consequently a more generic method to represent and detect deformable structures was investigated: active contour models. Together with an appropriate preprocessing which was developed in this research, active contours outperform the previously investigated approach, while at the same time being open for adaptions and extensions.

Approach	Correct contours	Correct sections
low-level	n/a	n/a
active contour	0/9	18/36 = 50%
fuzzy snake	8/9	35/36 = 97%

Table 5.11: Correct segmentation rates of different approaches.

The representational power of active contours however proved to be limited. Where a more detailed piece-wise description of an object is required, the representation of the

novel fuzzy active contours is superior. More complex contours or objects with a high degree of variations and uncertainties were successfully represented and segmented by the fuzzy active contour. Table 5.11 summarizes the performance of the investigated approaches.

Various parts of the traditional image processing solutions presented in this section were published in [4, 170–172]. The active contour approach was developed in [1, 4, 170–174]. The fuzzy active contour applied to the vocal tract X-rax images was published in [4].

Chapter 6

Conclusions

This thesis has presented both the theoretical basis and empirical analysis of a novel solution for segmenting and tracking of anatomical objects in 2D medical image sequences. An original combination of active contour models and fuzzy logic is employed to represent and integrate uncertain knowledge within the segmentation process. Experimental results on both synthetic and medical images validate the overall feasibility of the framework.

As a basis for the development of a new approach this research has thoroughly investigated the characteristics of medical image sequences both theoretically and through the design of clinical and scientific applications.

The investigations and developments were guided by a separation of the subject into image formation and object characteristics which correspond to a separation into low-level and high-level processing stages. It has been verified that the challenging properties of medical image sequences can be formulated as a set of variations and uncertainties in image formation and object characteristics. The observation that uncertainties play a major role in both low-level and high-level stages motivated the development of the fuzzy active contour, which provides a framework for integrating uncertain knowledge at all stages.

Existing image processing operators are also accounted for by the novel approach in defining an interface for their integration into the fuzzy framework. Traditional image processing solutions developed in this research suggested the combination of complementary operators to handle images of poor quality. For the detection and tracking of deformable objects in particular active contours were successfully extended with combined image processing operators.

Deferred decision making employed in this integration also became a key philosophy in the fuzzy active contour. Many existing approaches perform crisp decisions at each processing stage. Pixels are classified into possible objects or background through thresholded low-level operators and subsequently are grouped or deleted based on higher-level constraints. Conversely extended active contours developed in this research defer the segmentation decision to the last processing stage. Low-level operators are not thresholded early but are summed up to obtain continuous image evidences. A dynamic-programming-based snake optimisation further delays the final constraint-based segmentation. While this approach proved to be beneficial in a challenging application, the integration of *a priori* knowledge had limitations, which were overcome through the development of a novel contour representation.

Unlike existing representations the fuzzy contour model affords piece-wise description of image and geometric features. Both lengths and properties of each segment can be specified through fuzzy terms. These terms can be represented either as linguistic variables or grammatical characters, which allows for an exploitation of uncertain verbal knowledge as well as more formally defined *a priori* information.

A full integration of the high-level contour description into all stages of the segmentation was achieved through the introduction of fuzzy rules bases, each describing the properties of one contour segment. Instead of algebraic external energy functions with obscure numerical parameters, the fuzzy active contour calculates fuzzy external energies from linguistic rule bases.

The research has led to publications in the areas of image processing [1–5, 173], medical image analysis [162, 163, 170, 174], medical engineering [161, 164, 172] and fuzzy

logic [160, 175]. Early results were presented in [176] and also published in [171].

Points for further research

- Other object constraints than shape could be integrated into the framework to increase the representational power. Examples are velocity and direction of contour segments, their orientation as well as the distance to other segments or to landmarks which are not part of the contour.

- Representation schemes could be investigated that allow for a description of branched and looped contour topologies.

- The fuzzy snake bears the potential to be used a tool to assess and identify/classify previously segmented contours. It could be investigated how this capability can be realised using

  - E_{snake} as a crisp quality/classification measure and

  - a fuzzy quality measure based on the fuzzy E_{ext}.

- Alternative appearances of a contour segment can be represented within its rule base. To specify alternatives on the multi-segment level of the contour description, the logical OR could be integrated into the description and into the algorithm that selects a local rule base.

Bibliography

[1] HÖWING, F., D. WERMSER, and L.S. DOOLEY: *Recognition and tracking of articulatory organs in X-ray image sequences.* IEE Electronics Letters, 32(5):444–445, February 1996.

[2] HÖWING, F., L.S. DOOLEY, and D. WERMSER: *Tracking of non-rigid articulatory organs in X-ray image sequences.* Computerized Medical Imaging and Graphics, 23(2):19–27, April 1999.

[3] HÖWING, F., D. WERMSER, and L.S. DOOLEY: *Fuzzy snakes.* In *Int. Conf. Image Procssing and its Applications*, pages 627–630, Dublin, 1997. IEE Press.

[4] HÖWING, F., L.S. DOOLEY, and D. WERMSER: *Fuzzy active contour model.* IEE Proceedings on Vision, Image and Signal Processing, 147(4):323–330, August 2000.

[5] HÖWING, F., H. BÜLOW, D. WERMSER, L.S. DOOLEY, and W. THOMA: *Automatic motion analysis of bones from MR sequences.* In *Int. Conf. Image Processing and its Applications*, pages 397–401, Manchester, July 1999.

[6] PENTLAND, ALEX P.: *Perceptual organization and the representation of natural form.* Artificial Intelligence, 28:293–331, 1986.

[7] WECHSLER, HARRY: *Computational Vision.* Academic Press, 1990.

[8] BALLARD, DANA H. and CHRISTOPHER M. BROWN: *Computer Vision.* Prentice-Hall, 1982.

[9] RADIG, BERND (editor): *Verarbeiten und Verstehen von Bildern.* Oldenbourg, München, 1993.

[10] PINZ, AXEL: *Bildverstehen.* Springer, 1994.

[11] KASTURI, RANGACHAR and RAMESH C. JAIN (editors): *Computer Vision: Principles.* IEEE Computer Society Press, 1991.

[12] KASTURI, RANGACHAR and RAMESH C. JAIN (editors): *Computer Vision: Advances and Applications.* IEEE Computer Society Press, 1991.

[13] ERNST, HARTMUT: *Einführung in die digitale Bildverarbeitung: Grundlagen und industrieller Einsatz mit zahlreichen Beispielen.* Franzis, München, 1991.

[14] HABERÄCKER, PETER: *Digitale Bildverarbeitung: Grundlagen und Anwendungen.* Hanser, München; Wien, 1988.

[15] PRATT, WILLIAM K.: *Digital Image Processing.* John Wiley & Sons, New York, 1991.

[16] SONKA, MILAN, VACLAV HLAVAC, and ROGER BOYLE: *Image Processing, Analysis and Machine Vision.* Chapman & Hall, London, 1993.

[17] WAHL, FRIEDRICH M.: *Digitale Bildsignalverarbeitung: Grundlagen, Verfahren, Beispiele.* Springer, 1984.

[18] ZAMPERONI, PIERO: *Methoden der digitalen Bildsignalverarbeitung.* Vieweg, Braunschweig; Wiesbaden, 1989.

[19] ABMAYR, WOLFGANG: *Einführung in die digitale Bildverarbeitung.* Teubner, Stuttgart, 1994.

[20] PITAS, IOANNIS: *Digital Image Processing Algorithms.* Prentice Hall, 1993.

[21] SCHALKOFF, ROBERT J.: *Digital Image Processing and Computer Vision.* Wiley, 1989.

[22] TIZHOOSH, H.R.: *Fuzzy-Bildverarbeitung.* Springer, 1998.

[23] LIEDTKE, CLAUS-E. and MANFRED ENDER: *Wissensbasierte Bildverarbeitung.* Springer, 1989.

[24] CHO, Z.H., JOIE P. JONES, and MANBIR SINGH: *Foundations of Medical Imaging.* Wiley, 1993.

[25] LARSEN, C.F.: *Cineradiography of the wrist.* In BRUNELLI and SAFFAR [177], pages 100–106.

[26] SAFFAR, PH., C. SOKOLOW, C. MATHOULIN, Y. MARTIN-BOUYER, and S. VERDEILLE: *Cinearthrography of the wrist in carpal instability.* In BRUNELLI and SAFFAR [177], pages 109–114.

[27] LINSCHEID, R.L. and J.H. DOBYNS: *Karpale Instabilitäten.* Orthopäde, 22:72–78, 1993.

[28] PEICHA, G., F.J. SIBERT, M. FELLINGER, W. GRECHENIG and G. SCHIPPINGER: *Lesions of the Scapholunate Ligaments in Acute Wrist Trauma — Arthroscopic Diagnosis and Minimally Invasive Treatment.* Knee Surgery, Sports Traumatology, Arthroscopy, 5:176–183, 1997.

[29] METOUI, M.: *Chaos-Konzepte: Schlüssel zum Verstehen der Artikulationsdynamik.* In *Internationaler Kongress Sprache - Therapie - Computer.* Universität Graz, Institut für Sprachwissenschaften, April 1994.

[30] METOUI, M.: *Software-Einsatz für Forschung und Lehre in der artikulatorischen Phonetik.* In DETTE, K. and D. HAUPT (editors): *Multimedia und Computeranwendungen in der Lehre.* Springer, 1992.

[31] METOUI, M.: *Coarticulation: Statistical feature analysis with computer-aided processing of X-ray tracings.* In *6. Expertentreffen des Arbeitskreises Pädagogische Software mit digitaler Sprachverarbeitung.* Technische Universität Braunschweig, 1993.

[32] METOUI, M.: *Phonetische Datenverarbeitung mit S.P.A.T. - Möglichkeiten, Folgen, Herausforderungen.* Forschungsmagazin der Universität Mainz, Sonderheft, pages 18–25, March 1994.

[33] LIPPERT, HERBERT: *Anatomie: Text und Atlas.* Urban & Schwarzenberg, München; Wien; Baltimore, 5. edition, 1989.

[34] KRUGER, R.P., J.R. TOWNES, D.L. HALL, and S.J. DWYER III: *Automated radiographic diagnosis via feature extraction and classification of cardiac size and*

shape descriptors. IEEE Transactions on Biomedical Enineering, 19:174–186, 1972.

[35] TASTO, M.: *Guided boundary detection for left ventricular volume measurement*. In *Proceedings First International Joint Conference on Patter Recognition*, pages 119–124, Washington, D.C., October 1973.

[36] HOFMAN, I., H. NIEMANN, and G. SAGERER: *Model based interpretation from the heart*. In GELSEMA, EDZARD S. and LAVEEN N. KANAL [178].

[37] OLSTAD, B.: *Algorithms in Image Processing*. PhD thesis, Norwegian Institute of Technology, 1992.

[38] RUECKERT, D. and P. BURGER: *Shape-Based Segmentation and Tracking in 4D Cardiac MR Images*, pages 34–52. In TROCCAZ, J. et al. [179], March 1997.

[39] OMMEREN, J. VAN, C.J. KOOIJMAN, R.J. VAN MEENEN, J.J. GERBRANDS, A.V.M.C.L. SCHULTE, and J.H.C. REIBER: *Artery detection and analysis in cine-angiograms*. In GELSEMA, EDZARD S. and LAVEEN N. KANAL [178], pages 331–343.

[40] BEIER, J., H. OSWALD and E. FLECK: *Segmentierung von Gefässverzweigungen aus angiographischen Röntgenbildern*. In PINZ, A. (editor): *Wissensbasierte Mustererkennung*, volume 49 of *Schriftenreihe der Österreichischen Computer Gesellschaft*, pages 16–22, Wien; München, 1989. Oldenbourg.

[41] BEIER, J., H. OSWALD and E. FLECK: *Erkennung und Quantifizierung von Koronarstenosen aus angiographischen Röntgenbildern*. In BUNKE, HORST et al. [180], pages 24–30.

[42] COPPINI, G., M. DEMI, and G. VALLI: *An artificial vision system for X-ray images of human coronary trees*. IEEE Transactions on Pattern Analysis and Maschine Intelligence, 15(2):156–162, February 1993.

[43] HOARE, F. and G. DE JAGER: *Model based segmentation of medical X-ray images*. In *Image and Video Processing II*, pages 368–375. SPIE, February 1994.

[44] FRÖDER, MANFRED: *Darstellung geringkontrastiger Objekte im menschlichen Schädel mit rechnergestützter Röntgenvideotechnik.* PhD thesis, Universität Erlangen-Nürnberg, Technische Fakultät, 1986.

[45] TOLXDORFF, THOMAS, GEORG VIEHÖVER, BERTOLD WEIN and HEINZ HANDELS: *Automatische Segmentierung der Zunge in einer Ultraschallbildsequenz und Pseudo-3D-Darstellung der Zungenbewegung.* In BUNKE, HORST et al. [180], pages 354–360.

[46] BERGER, M.-O., G. MOZELLE, and Y. LAPRIE: *Cooperation of active contours and optical flow for tongue tracking in X-ray motion pictures.* In *Scandinavian Conf. On Image Analysis*, pages 913–920, 1995.

[47] PHAM, DZUNG L., CHENYANG XU, and JERRY L. PRINCE: *A survey of current methods in medical image segmentation.* Technical Report JHU/ECE 99-01, National Institute of Aging, Baltimore, 1998.

[48] BEZDEK, J.C., L.O. HALL, and L.P. CLARKE: *Review of MR image segmentation techniques using pattern recognition.* Med. Phys., 20:1033–1048, 1993.

[49] STAIB, L.H., A. CHAKRABORTY, and J.S. DUNCAN: *An integrated approach for locating neuroanatomical structure from MRI.* International Journal of Pattern Recognition and Artificial Intelligence, 11(8):1247–1269, 1997.

[50] BINFORD, THOMAS O.: *Survey of model-based image analysis systems.* International Journal of Robotics Research, 1(1):18–64, 1982.

[51] GERBRANDS, JAN JACOB: *Segmentation of noisy images.* PhD thesis, Technical University of Delft, 1988.

[52] BLAKE, A. and A. YUILLE (editors): *Active Vision.* MIT Press, Cambridge, Mass., 1992.

[53] MCINERNEY, T. and D. TERZOPOUOS: *Deformable models in medical image analysis: A survey.* Medical Image Analysis, 1(2):91–108, 1996.

[54] AYACHE, N.: *Medical image analysis and simulation.* In *Lecture Notes in Computer Science 1345*, pages 4–17. Springer, 1997.

[55] SINGH, A., D. GOLDGOF, and D. TERZOPOULOS: *Deformable Models in Medical Image Analysis*. IEEE Press, 1998.

[56] LI, H.D., M. KALLERGI, L.P. CLARKE, V.K. JAIN, and R.A. CLARK: *Markov random field for tumor detection in digital mammography*. IEEE Transactions on Medical Imaging, 14:565–576, 1995.

[57] LEE, C., S. HUN, T.A. KETTER, and M. UNSER: *Unsupervised connectivity-based thresholding segmentation of midsaggital brain MR images*. Comput. Biol. Med., 28:309–338, 1998.

[58] POHLMANN, S., K.A. POWELL, N.A. OBUCHOWSKI, W.A. CHILCOTE, and S. GRUNDFEST-BRONIATOWSKI: *Quantitative classification of breast tumors in digitzed mammograms*. Med. Phys., 23:1337–1345, 1996.

[59] KAPUR, T., W.E.L. GRIMSON, R. KIKINIS, and W.M. WELLS: *Enhanced spatial priors for segmentation of magnetic resonance imagery*. In *Proc. Int. Conf. Med. Im. Comput. Comp. Assist. Interven. (MICCAI98)*, pages 457–468, 1998.

[60] WELLS, W.M., W.E.L. GRIMSON, R. KIKINS, and F.A. JOLESZ: *Adaptive segmentation of MRI data*. IEEE Transactions on Medical Imaging, 15:429–442, 1996.

[61] REDDICK, W.E., J.O. GLASS, E.N. COOK, T.D. ELKIN, and R.J. DEATON: *Automated segmentation and classification of multispectral magnetic resonance images of brain using artificial neural networks*. IEEE Transactions on Medical Imaging, 16:911–918, 1997.

[62] RAJAPASKE, J.C., J.N. GIEDD, and J.L. RAPOPORT: *Statistical approach to segmentation of single-channel cerebral MR images*. IEEE Transactions on Medical Imaging, 16:176–186, 1997.

[63] HELD, K., E.R. KOPS, B.J. KRAUSE, W.M. WELLS, and R. KIKINIS: *Markov random field segmentation of brain MR images*. IEEE Transactions on Medical Imaging, 16(6), 1997.

[64] GEMAN, S. and D. GEMAN: *Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 6:721–741, 1984.

[65] HALL, L.O., A.M. BENSAID, L.P. CLARKE, R.P. VELTHUIZEN, M.S. SIL-BIGER, and J.C. BEZDEK: *A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain*. IEEE Transactions on Neural Networks, 3:672–682, 1992.

[66] CHRISTENSEN, G.E., S.C. JOSHI, and M.I. MILLER: *Volumetric transformation of brain anatomy*. IEEE Transactions on Medical Imaging, 16:864–877, 1997.

[67] THOMPSON, P. and A.W. TOGA: *Detection, visualization and animation of abnormal anatomic structure with a probabilistic brain atlas based on random vector field transformations*. Med. Im. Anal., 1:271–294, 1997.

[68] VOGT, M., M. SOMMERAU, G. MAMIER and P. LEVI: *Lokalisierung, Verfolgung und Modellierung von Lippen zur Audio-Visuellen Spracherkennung*, pages 464–471. In PAULUS, E. and F.M. WAHL [181], 1997.

[69] PFITZNER, K. and H. STRECKER: *XRAY – An Experimental Configuration Expert System for Automatic X-ray Inspection*. In PAULUS, E. (editor): *Mustererkennung 1987: 9. DAGM-Symposium*, pages 315–319. Dt. Arbeitsgemeinschaft für Mustererkennung, Springer, 1987.

[70] BÖHM, MICHAEL: *Entwicklung von Verfahren für die Verarbeitung und Analyse von Röntgenbildsequenzen auf Basis von Modellen der beobachteten Vorgänge*. PhD thesis, Universität Hamburg, Fachbereich Informatik, February 1985.

[71] GLOMBIZZA, G., M. MAKABE and H.P. MEINZER: *Gekrümmte Illusionskantenmodelle und Ihre Anwendung in der Bildverarbeitung*. In SAGERER, G., S. POSCH and F. KUMMERT (editors): *Mustererkennung*, pages 403–410. 1995.

[72] OLSTAD, B.: *Automatic wall motion detection in the left vertricle using ultrasonic images*. In *Proceedings of SPIE/SPSE: Electronic imaging. Science & Technology*, San Jose, 1991.

[73] CHEN, C.W., T.S. HUANG, and M. ARROT: *Modeling, analysis, and visualization of left ventricle shape and motion by hierarchical decomposition.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 16:342–356, 1994.

[74] KASS, MICHAEL, ANDREW WITKIN, and DEMETRI TERZOPOULOS: *Snakes: Active contour models.* In *Proceedings of International Conference on Computer Vision*, pages 259–268, London, June 1987.

[75] YUILLE, A.L., D.S. COHEN, and P.W. HALLINAN: *Feature extraction from faces using deformable templates.* Proc. Comp. Vision Pattern Recog., pages 104–109, 1989.

[76] LIPSON, P., A.L. YUILLE, D. O'KEEFFE, J. CAVANAUGH, J. TAAFFE, and D. ROSENTHAL: *Deformable templates for feature extraction from medical images.* In FAUGERAS, OLIVIER (editor): *Computer Vision — ECCV 90: First European Conference on Computer Vision*, pages 413–417. Springer, 1990.

[77] JAIN, A.K., Y. ZONG, and S. LAKSHMANAN: *Object matching using deformable templates.* IEEE Transactions on Pattern Analysis and Mashine Intelligence, 18(3):267–278, 1996.

[78] STAIB, L.H. and J.S. DUNCAN: *Boundary finding with parametrically deformable models.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(11):1061–1075, 1992.

[79] STAIB, L.H. and J.S. DUNCAN: *Model-based deformable surface finding for medical images.* IEEE Transactions on Medical Imaging, 15(5):720–731, 1996.

[80] SZEKELY, G., A. KELEMEN, CH. BRECHBUHLER, and G. GERIG: *Segmentation of 2-D and 3-D objects from MRI volume data using constrained elastic deformations of flexible Fourier surface models.* Medical Image Analysis, 1(1):19–34, 1996.

[81] COOTES, T.F., C.J. TAYLOR, D.H. COOPER, and J. GRAHAM: *Active shape models — their training and application.* Computer Vision and Image Understanding, 61(1):38–59, January 1995.

[82] MAHMOODI, S., B.S. SHARIF, E.G. CHESTER, J.P. OWEN, and R.E.J. LEE: *Automated vision system for skeletal age assessment using knowledge based techniques*. In *Image Procesing and its Applications*, pages 809–813, 1997.

[83] MA, Y. and D. C. WILKINS: *Induction of uncertain rules and the sociopathicity property in Dempster-Shafer theory*. In KRUSE, R. and P. SIEGEL (editors): *Europ. Conf. On Symbolic and Quantitative Approaches for Uncertainty*, pages 238–245. Springer, 1991.

[84] LIANG, Z.: *Tissue classification and segmentation of MR images*. IEEE Eng. Med. Biol., pages 81–85, 1993.

[85] SZELISKI, R.: *Bayesian modelling of uncertainty in low-level vision*. International Journal of Computer Vision, pages 271–301, 1990.

[86] TERZOPOULOS, D. and R. SZELISKI: *Tracking with Kalman Snakes*, pages 3–20. In BLAKE, A. and A. YUILLE [52], 1992.

[87] WILSON, A.G.: *Statistical Models for Shapes and Deformations*. PhD thesis, Duke University, 1995.

[88] FRITSCH, D.S., S.M. PIZER, L. YU, and E.L. CHANEY: *Segmentation of medical image objects using deformable shape loci*. In DUNCAN, J. and G. GINDI (editors): *Lecture Notes in Computer Science 1230*, pages 127–140. Springer, 1997.

[89] PIZER, S.M., D.S. FRITSCH, P. YUSHKEVICH, V. JOHNSON, and E.L. CHANEY: *Segmentation, registration, and measurement of shape variation via image object shape*. IEEE Transactions on Medical Imaging, 18:851–865, 1999.

[90] ISARD, M. and A. BLAKE: *Contour tracking by stochastic propagation of conditional density*. In BUXTON, B. and R. CIPOLLA (editors): *Computer Vision - ECCV '96*, pages 343–356. Springer, 1996.

[91] BLAKE, A., B. BASCLE, M. ISARD, and J.P. MACCORMICK: *Statistical models of visual shape and motion*. Phil. Trans. Royal Soc. London, A-356:1283–1302, 1998.

[92] KAPUR, T.: *Model Based Three Dimensional Medical Image Segmentation.* PhD thesis, AI Lab, MIT, 1999.

[93] DeCARLO, D. and D. METAXAS: *Blended deformable models.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(4):443–448, 1996.

[94] NASTAR, C. and N. AYACHE: *Frequency-based nonrigid motion analysis: Application to four dimensional medical images.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(11):1067–1079, 1996.

[95] ARNBORG, S.: *A survey of Bayesian data mining.* Technical Report 1100-3154, Swedish Institute of Computer Science, 2001.

[96] JAIN, A.K., R.P.W. DUIN, and J. MAO: *Statistical pattern recognition: A review.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(1):4–37, 2000.

[97] LAM, K.-M. and H. YAN: *Fast algorithm for locating head boundaries.* Journal of Electronic Imaging, 3(4):351–359, 1994.

[98] PARVIN, B., S. VISWANATHAN, and U. DAHMEN: *Tracking of convex objects.* In *Int Symp Computer Vision*, 1995.

[99] PARVIN, B., C. PENG, W. JOHNSTON, and M. MAESTRE: *Tracking of tubular molecules for scientific applications.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(8):800–805, 1995.

[100] LATOMBE, B., P. LADRET, F. GRANADA, and P. VILLEMAIN: *An original active contour algorithm applied to snow avalanches.* In *Int. Conf. Image Processing and its Applications*, pages 404–408, 1997.

[101] INGLIS, I.M. and A.J. GRAY: *Active contour models for segmenting hyphal images.* In *Int. Conf. Image Processing and its Applications*, pages 520–524, 1997.

[102] BERGER, M.-O.: *Snake growing.* In FAUGERAS, OLIVIER (editor): *Computer Vision — ECCV 90: First European Conference on Computer Vision*, pages 570–572. Springer, 1990.

[103] AYACHE, N., I. COHEN, and I. HERLIN: *Medical image tracking.* In BLAKE, A. and A. YUILLE [52], pages 285–302.

[104] STORVIK, GEIR: *A bayesian approach to dynamic contours through stochastic sampling and simulated annealing.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 16(10):976–986, October 1994.

[105] DAVIS: *Multiple energy function active contours applied to CT and MR images.* In *Int. Conf. Image Processing and its Applications*, pages 114–118, 1995.

[106] LARSEN, O.V., P. RADEVA, and E. MARTI: *Guidelines for choosing optimal parameters of elasticity for snakes.* In HLAVAC, V. and R. SARA [182], pages 106–113.

[107] KLEIN, A.K., F. LEE, and A.A. AMINI: *Quantitative coronary angiography with deformable spline models.* IEEE Transactions on Medical Imaging, 16:468–482, 1997.

[108] AMINI, AMIR A., SAEID TEHRANI, and TERRY E. WEYMOUTH: *Using dynamic programming for minimizing the energy of active contours in the presence of hard constraints.* In *Proceedings of the 2nd International Conference on Computer Vision*, pages 95–99, 1988.

[109] WILLIAMS, DONNA J. and MUBARAK SHAH: *A fast algorithm for active contours and curvature estimation.* Computer Vision, Graphics and Image Processing, 55(1):14–26, January 1992.

[110] RADEVA, R. and E. MARTI: *An Improved Model of Snakes for Model-Based Segmentation*, pages 515–520. In HLAVAC, V. and R. SARA [182], September 1995.

[111] MILLER, J.V., D.E. BREENAND W.E. LORENSEN, R.M. O'BARA, and M.J. WOZNY: *Geometrically deformed models: A method for extracting closed geometric models from volumetric data.* In SEDERBERG, T.W. (editor): *Proc. SIGGRAPH'91*, pages 217–226. ACM, 1991.

[112] COHEN, L.D.: *Note on active contour models and balloons*. Computer Vision, Graphics, and Image Processing: Image Understanding, 53(2):211–218, 1990.

[113] COHEN, L.D. and I. COHEN: *Finite-element methods for active contour models and balloons for 2D and 3D images*. IEEE Transactions on Patterns Analysis and Machine Intelligence, 15(11):1131–1147, 1993.

[114] IVINS, J. and J. PORILL: *Statistical snakes: Active region models*. In *British Machine Vision Conference*, pages 377–386, 1994.

[115] SCHNABEL, J.A. and S.R. ARRIDGE: *Multi-scale shape description in medical imaging*. In *Conf. Medical Image Understanding and Analysis*, pages 141–144, Oxford, 1997.

[116] BASCLE, B. and R. DERICHE: *Region tracking through image sequences*. In *Int. Conf. On Computer Vision*, pages 302–307. IEEE Computer Society Press, 1995.

[117] GRIMSON, W.E.L.: *Object Recognition by Computer: The Role of Geometric Constraints*. MIT Press, 1990.

[118] GUNN, S.R. and M.S. NIXON: *Improving Snake Performance via a Dual Active Contour*, pages 600–605. In HLAVAC, V. and R. SARA [182], September 1995.

[119] GUNN, S.R. and M.S. NIXON: *A robust snake implementation: A dual active contour*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(1):63–68, 1997.

[120] CASELLES, V., F. CATTE, T. COLL, and F. DIBES: *A geometric model for active contours in image processing*. Numerische Mathematik, 66:1–31, 1993.

[121] CASELLES, V., R. KIMMEL, and G. SAPIRO: *Geodesic active contours*. In *Int. Conf. Computer Vision*, pages 694–699, 1995.

[122] KICHENASSAMY, S., A. KUMAR, P. OLVER, A. TANNENBAUM, and A. YEZZI: *Gradient flows and geometric active contour models*. In *Int. Conf. Computer Vision*, pages 810–815, 1995.

[123] OLSTAD, B.: *Active contours with grammatical descriptions.* In *Proceedings of 6th International Conference on Image Analysis and Processing,* Como, Italy, September 1991.

[124] OLSTAD, B. and A.H. TORP: *Encoding of a priori information in active contour models.* Pattern Analysis and Machine Intelligence, 18(9):863–872, 1996.

[125] SEDGEWICK, ROBERT: *Algorithmen in C.* Addison-Wesley, 1992.

[126] BÄSSMANN, HENNING and PHILLIP W. BESSLICH: *Konturorientierte Verfahren in der digitalen Bildverarbeitung.* Springer, 1989.

[127] SÜSSE, H.: *Liniendetektion mittels zeiteffektiver Dynamischer Programmierung.* In GROSSKOPF, R.E. [183], pages 256–266.

[128] GEIGER, DAVI, ALOK GUPTA, LUIZ A. COSTA, and JOHN VLONTZOS: *Dynamic programming for detecting, tracking, and matching deformable contours.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(3):294–302, March 1995.

[129] ZADEH, LOFTI ASKER: *Fuzzy sets.* Information and Control, 9(8):338–353, August 1965.

[130] ZADEH, LOFTI ASKER: *The concept of a linguistic variable and its application to approximate reasoning – I.* Information Sciences, 10(8):220–222, August 1975.

[131] ZIMMERMANN, H.-J.: *Fuzzy Set Theory and its Applications.* Kluwer, 1991.

[132] PAL, S.K. and D.K.D. MAJUMDER: *Fuzzy Mathematical Approach to Pattern Recognition.* Wiley, 1986.

[133] BEZDEK, J.C. and S.K. PAL: *Fuzzy Methods for Pattern Recognition.* IEEE Press, New York, 1992.

[134] BANDEMER, H. and S. GOTTWALD: *Einführung in Fuzzy-Methoden.* Akademie, Berlin, 1993.

[135] DUBOIS, D., H. PRADES, and R.R. YAGER (editors): *Fuzzy Information Engineering.* Wiley, 1997.

[136] GOTTWALD, S.: *Fuzzy Sets and Fuzzy Logic*. Vieweg, Wiesbaden, 1993.

[137] MARTIN, T.P. and A.L. RALESCU (editors): *Fuzzy Logic in Artificial Intelligence*. Springer, 1997.

[138] SCHULTE, U.: *Einführung in Fuzzy Logik*. Franzis, München, 1993.

[139] TILLI, T.: *Mustererkennung und Fuzzy Logik*. Franzis, München, 1993.

[140] TRAEGER, D.: *Einführung in die Fuzzy-Logik*. Teubner, Stuttgart, 1993.

[141] BEZDEK, J.C., L.O. HALL, M. CLARK, D. GOLDOF, and L.P. CLARKE: *Segmenting Medical Images with Fuzzy Models: An Update*, chapter 4, pages 69–92. In DUBOIS, D. et al. [135], 1997.

[142] CLARK, M.C., L.O. HALL, D.B. GOLDGOF, and M.S. SILBIGER: *Using Fuzzy Information in Knowledge Guided Segmentation of Brain Tumors*, pages 167–181. In MARTIN, T.P. and A.L. RALESCU [137], 1997.

[143] PAL, S.K. and R.A. KING: *On edge detection of X-ray images using fuzzy sets*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 5(1):69–77, 1983.

[144] LAW, T., H. ITOH, and H. SEKI: *Image filtering, edge detection, and edge tracing using fuzzy reasoning*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(5):481–491, 1996.

[145] HO, K.H.L. and N. OHNISHI: *FEDGE - Fuzzy Edge Detection by Fuzzy Categorization and Classification*, pages 182–196. In MARTIN, T.P. and A.L. RALESCU [137], 1997.

[146] KNOPF, G.K. and M.M. GUPTA: *Fuzzy uncertainty measures in image processing*. Journal of Electronic Imaging, 3(2):142–153, 1994.

[147] PATHAK, A. and S.K. PAL: *Fuzzy grammars in syntactic recognition of skeletal maturity from X-rays*. IEEE Transactions on Systems, Man, and Cybernetics, 16(5):657–667, 1986.

[148] ASVELD, P.R.J.: *A bibliography on fuzzy automata, grammars and languages.* In *Bulletin of the European Association for Theoretical Computer Science*, pages 187–196, 1996.

[149] PARIZEAU, M. and R. PLAMONDON: *A fuzzy-syntactic approach to allograph modeling for cursive script recognition.* IEEE Transactions on Pattern Analysis and Maschine Intelligence, 17(7):702–712, 1995.

[150] RALESCU, A.L. and J.G. SHANAHAN: *Structure Cognition from Images*, pages 225–239. In MARTIN, T.P. and A.L. RALESCU [137], 1997.

[151] JONES, T.N. and D.N. METAXAS: *Segmentation using deformable models with affinity-based localisation.* In TROCCAZ, J. et al. [179], pages 53–62.

[152] BOLLMANN, M., R. HOISCHEN, and B. MERTSCHING: *Integration of static and dynamic scene features.* In PAULUS, E. and F.M. WAHL [181], pages 483–490.

[153] VOEGTLE, T. and R. JÄGER: *Detektion bewegter starrer Objekte in Realwelt-Bildfolgen mit Hilfe von Differenzbildern.* In GROSSKOPF, R.E. [183], pages 675–681.

[154] ARMBRECHT, T.: *Definition und Implementierung von Form-Constraints für eine Konturerkenung mit Mitteln der Fuzzy-Logic.* Master's thesis, Fachhochschule Braunschweig/Wolfenbüttel, Fachbereich Elektrotechnik, December 1997.

[155] KACAR, T.: *Demonstration der Funktionsweise von Fuzzy-Snake Verfahren gegenüber herkömmlichen Aktiven Konturen anhand exemplarischer Testkonturen.* Master's thesis, Fachhochschule Braunschweig/Wolfenbüttel, June 1998.

[156] ROTHKAMM, DIRK: *Entwicklung und Software-Implementation eines Verfahrens zur Lokalisierung der Region of Interest in MRT-Bildsequenzen.* Master's thesis, Fachhochschule Braunschweig/Wolfenbüttel, FB Elektrotechnik, February 1998.

[157] CHEN, Q., M. DEFRISE and F. DECONINCK: *Symmetric Phase-Only Matched Filtering of Fourier-Mellin Transforms for Image Registration and Recognition.* IEEE Pattern Analysis and Machine Intelligence, 16(12):1156–1168, 1994.

[158] FRÜHAUF, JENS: *Automatische Auswahl geeigneter Schichten in MRT-Aufnahmen der Handwurzelknochen auf der Basis des Fourier-Mellin Verfahrens.* Master's thesis, Fachhochschule Braunschweig/Wolfenbüttel, FB Elektrotechnik, January 1999.

[159] AMOROS, MARIA JESUS SOGORB: *Automatische Segmentierung und Identifikation der Handwurzelknochen in MRT-Aufnahmen.* Master's thesis, Fachhochschule Braunschweig/Wolfenbüttel, FB Elektrotechnik, October 1998.

[160] HÖWING, F., L.S. DOOLEY, and D. WERMSER: *Linguistic contour modelling through a fuzzy active contour.* In MOHAMMADIAN, M. (editor): *New Frontiers in Computational Intelligence and its Applications,* pages 271–279. IOS Press, Amsterdam, 2000.

[161] DOOLEY, L.S., F. HÖWING, and D. WERMSER: *Automatic analysis of MR sequences for the diagnosis of ligament lesions.* In *European Medical and Biological Engineering Conference,* pages 934–935, Vienna, November 1999. IEE Press.

[162] HÖWING, F., H. BÜLOW, D. WERMSER, L.S. DOOLEY and W. THOMA: *Analyse von Knochenbewegungen in Folgen von MRT-Aufnahmen.* In *Bildverarbeitung für die Medizin,* pages 317–321. Springer, March 1999.

[163] HÖWING, F., D. WERMSER and L.S. DOOLEY: *Linguistische Modellierung zur Erkennung anatomischer Objekte.* In *Bildverarbeitung für die Medizin,* pages 357–361. Springer, March 1999.

[164] THOMA, W., M. GÖTZ, C.-H. FÜRST, F. HÖWING, H. BÜLOW and D. WERMSER: *MRT-Bilderkennungssystem zur automatischen Analyse komplexer Knochenbewegungen.* In *Deutscher Orthopädenkongreß, Wiesbaden,* September 1998. Scientific video.

[165] MEYER, TORSTEN: *Erarbeitung und Untersuchung von kantenorientierten Verfahren zur Vermessung der Elemente des Vokaltrakts in Röntgenbildern.* Master's thesis, Fachhochschule Braunschweig/Wolfenbüttel, FB Elektrotechnik, December 1995.

[166] MATTHIAS, PETER: *Positionsbestimmung starrer Objekte in Röntgenbildsequenzen durch Template Matching in der Bildebene*. Master's thesis, Fachhochschule Braunschweig/Wolfenbüttel, FB Elektrotechnik, June 1997.

[167] GOSHTASBY, ARDESHIR: *Template Matching in Rotated Images*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 7(3):338–344, May 1985.

[168] ENDERS, ROLF: *Bewegungskompensation in Röntgenbildsequenzen durch rotationsinvariantes Template Matching*. Master's thesis, Fachhochschule Braunschweig/Wolfenbüttel, FB Informatik, April 1997.

[169] STOLPMANN, A. and L.S. DOOLEY: *A Texture Classification System Using Statistical and Soft-Computing Methods*. In *Proceedings of the 1st Infotech Oulu Workshop on Texture Analysis in Machine Vision*, Oulu, Finland, 1999.

[170] HÖWING, F., D. WERMSER and L.S. DOOLEY: *Entwicklung Automatischer Verfahren zur Kennzeichnung und Vermessung von verformbaren Objekten in Bildsequenzen*. In *Bildverarbeitung für die Medizin*, Heidelberg, March 1999. Springer. Scientific video.

[171] HÖWING, F.: *Automatic Identification and Measurement of X-Ray Sequences Involving the Articulatory Tract*. MPhil to PhD transfer document, University of Glamorgan, March 1996.

[172] WERMSER, D. and F. HÖWING: *Automatische Kennzeichnung und Vermessung in Röntgenbildsequenzen*. Biomedical Journal, (45):17–19, January 1996.

[173] HÖWING, F., L.S. DOOLEY and D. WERMSER: *Tongue Tracking in Medical X-Ray Sequences*. In *Int. Conf. Image Processing and its Applications*, pages 494–497, Manchester, July 1999.

[174] HÖWING, F., L.S. DOOLEY and D. WERMSER: *Zungenverfolgung in medizinischen Röntgenbildsequenzen*. In *Bildverarbeitung für die Medizin*, pages 377–381. Springer, March 1999.

[175] HÖWING, F., L.S. DOOLEY, and D. WERMSER: *Linguistic contour modelling through fuzzy snakes.* In *Int. Conf. Computational Intelligence for Modelling Control and Automation*, pages 384–389, Vienna, February 1999. IOS Press.

[176] WERMSER, D. and F. HÖWING: *Automatische Vermessung in Röntgenbildsequenzen: Lautbildung im Vokaltrakt.* Exhibition and technical report, CeBIT'95, Hannover, Fachhochschule Braunschweig/Wolfenbüttel, 1995.

[177] BRUNELLI and SAFFAR (editors): *Wrist Imaging.* 1992.

[178] GELSEMA, EDZARD S. and LAVEEN N. KANAL (editors): *Pattern Recognition in Practice II*, Amsterdam; New York, 1986. Elsevier.

[179] TROCCAZ, J., E. GRIMSON, and R. MÖSGES (editors): *VCRMed-MRCAS97, First Joint Conf. Computer Vision, Virtual Reality and Robotics in Medicine and Medical Robotics and Computer-Assisted Surgery.* Springer, March 1997.

[180] BUNKE, HORST, OLAF KÜBLER and PETER STUCKI (editors): *Mustererkennung 1988: 10. DAGM-Symposium.* Dt. Arbeitsgemeinschaft für Mustererkennung, Springer, 1988.

[181] PAULUS, E. and F.M. WAHL (editors): *Mustererkennung*, 1997.

[182] HLAVAC, V. and R. SARA (editors): *CAIP'95 - Computer Analysis of Images and Patterns.* Springer, September 1995.

[183] GROSSKOPF, R.E. (editor): *Mustererkennung 1990: 12. DAGM-Symposium.* Dt. Arbeitsgemeinschaft für Mustererkennung, Springer, 1990.

[184] BALZERT, H.: *Die Entwicklung von Software-Systemen: Prinzipien, Methoden, Sprachen, Werkzeuge.* Bibliographisches Institut, Mannheim, Wien, Zürich, 1982.

[185] OTT, H.-J.: *Software-Systementwicklung: Praxisorientierte Verfahren und Methoden.* Hanser, Wien, München, 1991.

[186] HERING, E.: *Software Engineering.* Vieweg, Braunschweig, Wiesbaden, 1992.

Appendix A

Structured Analysis and Design Technique

Throughout the research project the *Structured Analysis and Design Technique* (SADT) has been used as a graphical method to describe system structure and to document functional units.

When applying SADT, a system is described in a structured and hierarchical way. In a top-down approach the system structure is first designed coarsely. The functional units are detailed then in a number of sub-diagrams, until the systems specification is complete.

SADT differentiates between activity and data diagrams. With activity diagrams an action is shown as a box while data flow follows arrows between the boxes. Data diagrams reverse this definition and are not used in this thesis.

Each side of the box may have arrow(s) of the following meanings:

- *left:* input data (I), will be transformed to output data;

- *top:* control data (C), influences or controls the transformation;

- *right:* output data (O), the result of the action;

- *bottom:* mechanism (M), the executing element or an aid to the action.

The letters in parenthesis identify the elements of the action's interface at a lower level. Each action has a number, which is used to identify its lower-level diagram. The top-level diagram is labelled *A0*, the diagram of action 1 in *A0* is called *A1*, the diagram of the third action in *A1* is marked *A13* and so on.

For more information on SADT consider [184–186].

Appendix B

Selected Publications

Articles in Books

- Höwing, F., Dooley L.S., Wermser, D.: *Linguistic Contour Modelling through a Fuzzy Active Contour.* Mohammadian, M. (Ed.), New Frontiers in Computational Intelligence and its Applications, invited paper, pp. 271–279, Amsterdam 1999

Journal Articles

- Höwing, F., Dooley L.S., Wermser, D.: *Fuzzy Active Contour Model.* IEE Proceedings on Vision, Image and Signal Processing, invited paper, 147(4):323-330, August 2000

- Höwing, F., Dooley L.S., Wermser, D.: *Tracking of non-rigid articulatory organs in X-ray image sequences.* Computerized Medical Imaging and Graphics. 23(2):19-27, April 1999

- Höwing, F., Wermser, D. and Dooley L.S.: *Recognition and Tracking of Articulatory Organs in X-ray Image Sequences.* IEE Electronics Letters, 32(5):444–445, February 1997

Articles in Conference Proceedings

- Höwing, F., Bülow, H., Wermser, D., Dooley, L.S. and Thoma, W.: *Automatic Motion Analysis of Bones from MR Sequences.* International Conference on Image Processing and its Applications, pp. 397–401, Manchester, July 1999

- Höwing, F., Wermser, D. and Dooley, L.S.: *Fuzzy Snakes.* International Conference on Image Processing and its Applications, pp. 627–630, Dublin, July 1997

Linguistic Contour Modelling
through a Fuzzy Active Contour

Frank Höwing
University of Glamorgan, School of Electronics, UK
also with
f.hoewing@fh-wolfenbuettel.de
Fachhochschule Braunschweig/Wolfenbüttel, FB E, Germany

Laurence S Dooley
Monash University, Gippsland Campus,
School of Computing and Information Technology,
Churchill, Victoria 3842, Australia
laurence.dooley@infotech.monash.edu.au

Diederich Wermser
Fachhochschule Braunschweig/Wolfenbüttel, FB E
Salzdahlumer Str. 46/48, 38302 Wolfenbüttel, Germany
wermser@fh-wolfenbuettel.de

Abstract. A new method for representing and tracking of object boundaries is presented, which allows for the integration of uncertain *a priori* knowledge into an active contour model.
The novel concept of fuzzy snakes is developed to allow for an intuitive specification of the properties of an object's boundary. This is achieved by introducing fuzzy energy functions and establishing a linguistic rule base, which describes each of the fuzzy snake's segments. Furthermore the approximate length of each contour segment may be specified to both improve the segmentation process and to reduce computational complexity.
Finally, a complex scene from a medical imaging sequence has been included to illustrate the performance of this fuzzy contouring technique.

1 Introduction

Active contours [1], or *snakes*, are a well known method for matching an object's contour model to features in an image. The approach, which uses a polygonal object representation (Fig. 1), is distinguished by its intrinsic ability to handle variations in the boundary that is to be detected. It is therefore capable of identifying and tracking deformable objects in image sequences.

For each image, the algorithm requires an initial polygon $P = (p_0, p_1, \ldots, p_{N-1})$ consisting of N vertices $p_i = (x_i, y_i)$, where x_i and y_i are the spatial co-ordinates of p_i. The detected boundary is represented by the polygon $Q = (q_0, q_1, \ldots, q_{N-1})$, with $q_i = (x_i, y_i)$. Each q_i is selected from a set of candidates $C_i = (c_{i,0}, c_{i,1}, \ldots, c_{i,M-1})$. In many applications the candidates $c_{i,j} = (x_j, y_j)$ are uniformly sampled along a search line normal to the initial polygon and intersecting p_i. Tracking of the contour is achieved by processing a sequence frame by frame and taking the resulting Q_t as the initial estimated contour P_{t+1} for the next frame.
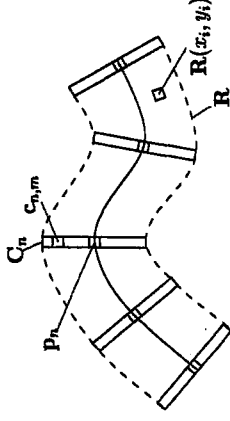
F. Höwing et al. / Linguistic Contour Modelling

Figure 1: Polygonal contour representation.

From this polygonal representation, it is possible to formulate an appropriate energy function E_{snake} for the object, which can subsequently be minimised in order to obtain the desired contour Q. This minimisation is achieved by selecting an optimal set of vertices from the candidates which have been sampled in a region R, around an initial contour.

Normally E_{snake} is decomposed into two components:

$$E_{snake}(Q) = \sum_{n=0}^{N-1} w_{int} E_{int}(q_n) + w_{ext} E_{ext}(q_n) \qquad (1)$$

E_{int}, the internal energy represents general contour properties such as stiffness, which are usually required to produce a smooth shape. The external energy E_{ext} is composed of E_{image} which guides the contour towards particular features in the image and E_{con} which allows for the integration of additional constraints (cf. equation (2)). Weighting parameters w_{int}, w_{ext}, w_{image}, and w_{con} control the relative influence of the energy components and are generally determined by a process of trial-and-error.

$$E_{ext}(q_n) = w_{image} E_{image}(q_n) + w_{con} E_{con}(q_n) \qquad (2)$$

The original active contour algorithm presented in [1] had some inherent computational problems in evaluating the energy function, which were subsequently solved by Amini et al. [2]. The energy minimisation of the snake was performed by a discrete dynamic programming algorithm which allowed for the integration of hard constraints, such as a minimum distance between the snake's vertices.

2 Active Contours with Multiple Segments

A fundamental limitation remained in terms of representing *a priori* knowledge concerning more complex objects which were to be detected, namely that all vertices of the active contour were characterised by the same local energy function, resulting in a single, global description of the object. In order to overcome this problem, Olstad [3,4] introduced a grammatical description of the snake's energy function. Fig. 2 shows an example of a contour, which can be described as a sequence of four different external energy functions E_{ext}^a, E_{ext}^b, E_{ext}^c and E_{ext}^d, represented by the respective terminals a, b, c and d. The grammatical expression describing such a segmented boundary would be a*b*c*d*, where * is the closure operator which allows parts of a pattern to be arbitrarily repeated. A pattern-matching algorithm is incorporated within the active contour's energy minimisation process which constrains the resulting contours to only those that comply with the grammatical expression.
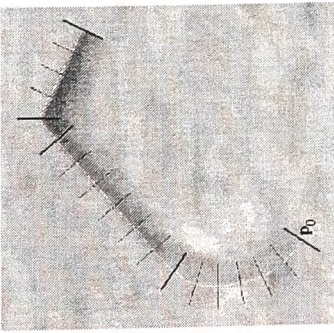
Figure 2: Example object. This synthetic boundary demonstrates a multi-segment object contour. The boundary segments have a different shape (local curvature) as well as a different appearance in the image (edge contrast), while all vertices within a segment share similar properties. The contour's vertices p_n are visualised by lines perpendicular to the contour. Thick lines illustrate the segment boundaries.

Three fundamental drawbacks in this algorithm can be identified:

1. The different energy functions do not intrinsically consider inexact a priori knowledge.

2. The length of a contour segment cannot be specified, although an approximate value may be known in advance.

3. The algorithm is computationally expensive, since the closure operation generates a large number of possible states in the finite-state-machine based pattern matching.

This paper focuses upon these specific disadvantages and introduces a new technique based upon fuzzy snakes [5] which will be shown to solve all three problems, the latter two having a common solution.

3 Fuzzy Segment Length

The fuzzy snake approach permits the specification of the number of subsequent vertices that share a common energy function, a parameter constraint which dramatically reduces the search space. A crisp length specification however, would not be able to consider uncertain information. A new method to specify the length of a snake segment by a fuzzy number is therefore presented.

3.1 Notation

Instead of specifying snake segments of arbitrary length as for example a*b*c*d*, the length of each segment can be given by the expression aaaaabbbbbcdddd. Since segment lengths are not known precisely, a different notation is introduced, where the length is expressed as a (discrete) fuzzy number [6] l, with mean value l_0 and spread s, defined by $\mu(l) > 0 \ \forall \ l \in \{l_0 - s, \ldots, l_0 + s\}$.

The contour of the example object in Fig. 2 can now be described by the expression $(6a)(6b)(1c)(4d)$. In the first segment, for a fuzzy length of $l_0 = l_a = 6$ and a spread

of $s = 2$, the segment specification $(6a)$ denotes a segment consisting of between 4 and 8 vertices, sharing a common property defined by an energy function E_{ext}^a.

Using the analogy of a grammatical description, the fuzzy snake can now be specified as an expression, using operators such as concatenation, logical AND and OR, together with a fuzzy length.

The absolute number of vertices depends on the sampling distance and is not a direct measure for the length of a segment. While the fuzzy snake algorithm uses the absolute length, the user-level contour description must allow for a length measure relative to the overall length of the contour. With the segment length expressed as a percentage, the above example can be written as $(0.4a)(0.4b)(0.05c)(0.15d)$.

3.2 Extended optimisation algorithm

An extension to the original dynamic programming algorithm [7] for global optimisation of the energy function is necessary to account for the variable length of snake segments.

The segment length is regarded as an additional constraint, where $l_z(n-1,k)$ is analogous to $S_z(n-1,k)$ and denotes the number of preceding vertices which would fall into state z if candidate $c_{n,m}$ was selected. Consequently, candidates which assist in constructing a chain of the specified length are favoured. Line 4 of Table 1 updates the length information. T_z in line 5 points to that predecessor of $c_{n,m}$ which would be optimal if the final state of vertex n was z.

1. **for** $n = 1 \ldots N - 1$
2.   **for** $m = 0 \ldots M - 1$
3. $S_z(n,m) = \min\limits_k \left[w_{int} E_{int}(c_{n,m}) + w_{ext} E_{ext}^a(c_{n,m}) + S_z(n-1,k) + 1 - \mu(l_z(n-1,k)) \right]$
4. $l_z(n,m) = l_z^k(n-1, k^{min}) + 1$
5. $T_z(n,m) = k^{min}$

Table 1: Extended dynamic programming algorithm.

An important additional advantage occurring from this approach is that computational complexity is reduced considerably. In the above example, an arbitrary segment length would generate $(N-1)^{(Z-1)} = 15^3 = 3375$ possible states of the contour, where Z is the number of segments. Exploiting the a priori knowledge concerning the approximate segment length reduces this number to 52, an approximate saving of nearly 98%. At a cursory glance this appears to be an inordinate improvement in performance, but it must be clearly stated however that the actual computational complexity is very much application-specific and depends on l_s, s and the complexities $O(E_{ext}^a)$, where $Z \in \{a, b, c, \ldots\}$. Nevertheless the integration of a priori information upon segment length into the active contour does have significant benefits.

4 Fuzzy Energy Functions

The second novel feature integrated into the proposed fuzzy snake is the fuzzy representation of energy functions. Using linguistic variables, this approach provides the active contour with an intuitive man-machine-interface, allowing uncertain knowledge to be exploited.

Linguistic Contour Modelling
through a Fuzzy Active Contour

Frank Höwing

University of Glamorgan, School of Electronics, UK

also with

Fachhochschule Braunschweig/Wolfenbüttel, FB E, Germany

f.hoewing@fh-wolfenbuettel.de

Laurence S Dooley

Monash University, Gippsland Campus,
School of Computing and Information Technology,
Churchill, Victoria 3842, Australia

laurence.dooley@infotech.monash.edu.au

Diederich Wermser

Fachhochschule Braunschweig/Wolfenbüttel, FB E,
Salzdahlumer Str. 46/48, 38302 Wolfenbüttel, Germany

wermser@fh-wolfenbuettel.de

Abstract. A new method for representing and tracking of object boundaries is presented, which allows for the integration of uncertain *a priori* knowledge into an active contour model.

The novel concept of fuzzy snakes is developed to allow for an intuitive specification of the properties of an object's boundary. This is achieved by introducing fuzzy energy functions and establishing a linguistic rule base, which describes each of the fuzzy snake's segments. Furthermore the approximate length of each contour segment may be specified to both improve the segmentation process and to reduce computational complexity.

Finally, a complex scene from a medical imaging sequence has been included to illustrate the performance of this fuzzy contouring technique.

1 Introduction

Active contours [1], or *snakes*, are a well known method for matching an object's contour model to features in an image. The approach, which uses a polygonal object representation (Fig. 1), is distinguished by its intrinsic ability to handle variations in the boundary that is to be detected. It is therefore capable of identifying and tracking deformable objects in image sequences.

For each image, the algorithm requires an initial polygon $P = (p_0, p_1, \ldots, p_{N-1})$ consisting of N vertices $p_i = (x_i, y_i)$, where x_i and y_i are the spatial co-ordinates of p_i. The detected boundary is represented by the polygon $Q = (q_0, q_1, \ldots, q_{N-1})$, with $q_i = (x_i, y_i)$. Each q_i is selected from a set of candidates $C_i = (c_{i,0}, c_{i,1}, \ldots, c_{i,M-1})$. In many applications the candidates $c_{i,j} = (x_j, y_j)$ are uniformly sampled along a search line normal to the initial polygon and intersecting p_i. Tracking of the contour is achieved by processing a sequence frame by frame and taking the resulting Q_t as the initial estimated contour P_{t+1} for the next frame.
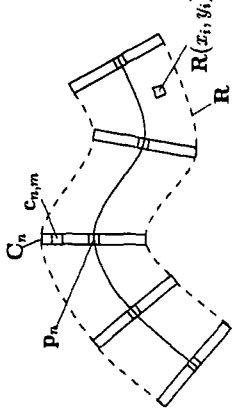
Figure 1: Polygonal contour representation.

From this polygonal representation, it is possible to formulate an appropriate energy function E_{snake} for the object, which can subsequently be minimised in order to obtain the desired contour Q. This minimisation is achieved by selecting an optimal set of vertices from the candidates which have been sampled in a region R, around an initial contour.

Normally E_{snake} is decomposed into two components:

$$E_{snake}(Q) = \sum_{n=0}^{N-1} w_{int} E_{int}(q_n) + w_{ext} E_{ext}(q_n) \qquad (1)$$

E_{int}, the internal energy represents general contour properties such as stiffness, which are usually required to produce a smooth shape. The external energy E_{ext} is composed of E_{image} which guides the contour towards particular features in the image and E_{con} which allows for the integration of additional constraints (cf. equation (2)). Weighting parameters w_{int}, w_{ext}, w_{image}, and w_{con} control the relative influence of the energy components and are generally determined by a process of trial-and-error.

$$E_{ext}(q_n) = w_{image} E_{image}(q_n) + w_{con} E_{con}(q_n) \qquad (2)$$

The original active contour algorithm presented in [1] had some inherent computational problems in evaluating the energy function, which were subsequently solved by Amini et al. [2]. The energy minimisation of the snake was performed by a discrete dynamic programming algorithm which allowed for the integration of hard constraints, such as a minimum distance between the snake's vertices.

2 Active Contours with Multiple Segments

A fundamental limitation remained in terms of representing *a priori* knowledge concerning more complex objects which were to be detected, namely that all vertices of the active contour were characterised by the same local energy function, resulting in a single, global description of the object. In order to overcome this problem, Olstad [3,4] introduced a grammatical description of the snake's energy function. Fig. 2 shows an example of a contour, which can be described as a sequence of four different external energy functions E_{ext}^a, E_{ext}^b, E_{ext}^c, E_{ext}^d and E_{ext}^d, represented by the respective terminals a, b, c and d. The grammatical expression describing such a segmented boundary would be a*b*c*d*, where * is the closure operator which allows parts of a pattern to be arbitrarily repeated. A pattern-matching algorithm is incorporated within the active contour's energy minimisation process which constraints the resulting contours to only those that comply with the grammatical expression.
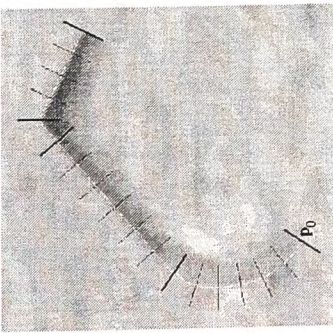
Figure 2: Example object. This synthetic boundary demonstrates a multi-segment object contour. The boundary segments have a different shape (local curvature) as well as a different appearance in the image (edge contrast), while all vertices within a segment share similar properties. The contour's vertices p_n are visualised by lines perpendicular to the contour. Thick lines illustrate the segment boundaries.

Three fundamental drawbacks in this algorithm can be identified:

1. The different energy functions do not intrinsically consider inexact *a priori* knowledge.

2. The length of a contour segment cannot be specified, although an approximate value may be known in advance.

3. The algorithm is computationally expensive, since the closure operation generates a large number of possible states in the finite-state-machine based pattern matching.

This paper focuses upon these specific disadvantages and introduces a new technique based upon fuzzy snakes [5] which will be shown to solve all three problems, the latter two having a common solution.

3 Fuzzy Segment Length

The fuzzy snake approach permits the specification of the number of subsequent vertices that share a common energy function, a parameter constraint which dramatically reduces the search space. A crisp length specification however, would not be able to consider uncertain information. A new method to specify the length of a snake segment by a fuzzy number is therefore presented.

3.1 Notation

Instead of specifying snake segments of arbitrary length as for example a*b*c*d*, the length of each segment can be given by the expression aaaaaabbbbbbbbcddd. Since segment lengths are not known precisely, a different notation is introduced, where the segment length is expressed as a (discrete) fuzzy number [6] l, with mean value l_0 and spread s, defined by $\mu(l) > 0 \ \forall \ l \in \{l_0 - s, \ldots, l_0 + s\}$.

The contour of the example object in Fig. 2 can now be described by the expression $(6a)(6b)(1c)(4d)$. In the first segment, for a fuzzy length of $l_0 = l_a = 6$ and a spread

of $s = 2$, the segment specification $(6a)$ denotes a segment consisting of between 4 and 8 vertices, sharing a common property defined by an energy function E_{ext}^a.

Using the analogy of a grammatical description, the fuzzy snake can now be specified as an expression, using operators such as concatenation, logical AND and OR, together with a fuzzy length.

The absolute number of vertices depends on the sampling distance and is not a direct measure for the *length* of a segment. While the fuzzy snake algorithm uses the absolute length, the user-level contour description must allow for a length measure relative to the overall length of the contour. With the segment length expressed as a percentage, the above example can be written as $(0.4a)(0.4b)(0.05c)(0.15d)$.

3.2 Extended optimisation algorithm

An extension to the original dynamic programming algorithm [7] for global optimisation of the energy function is necessary to account for the *variable* length of snake segments.

The segment length is regarded as an additional constraint, where $l_z(n-1, k)$ is analogous to $S_z(n-1, k)$ and denotes the number of preceding vertices which would fall into state z if candidate $c_{n,m}$ was selected. Consequently, candidates which assist in constructing a chain of the specified length are favoured. Line 4 of Table 1 updates the length information. T_z in line 5 points to that predecessor of $c_{n,m}$ which would be optimal if the final state of vertex n was z.

1. **for** $n = 1 \ldots N - 1$
2.   **for** $m = 0 \ldots M - 1$
3. $S_z(n,m) = \min\limits_{k} \left[w_{int} E_{int}(c_{n,m}) + w_{ext} E_{ext}^z(c_{n,m}) + S_z(n-1, k) + 1 - \mu(l_z(n-1, k)) \right]$
4. $l_z(n,m) = l_z(n-1, k^{min}) + 1$
5. $T_z(n,m) = k^{min}$

Table 1: Extended dynamic programming algorithm.

An important additional advantage occurring from this approach is that computational complexity is reduced considerably. In the above example, an arbitrary segment length would generate $(N-1)^{(Z-1)} = 15^3 = 3375$ possible states of the contour, where Z is the number of segments. Exploiting the *a priori* knowledge concerning the approximate segment length reduces this number to 52, an approximate saving of nearly 98%. At a cursory glance this appears to be an inordinate improvement in performance, but it must be clearly stated however that the actual computational complexity is very much application-specific and depends on l, s and the complexities $O(E_{ext}^z)$, where $Z \in \{a, b, c, \ldots\}$. Nevertheless the integration of *a priori* information upon segment length into the active contour does have significant benefits.

4 Fuzzy Energy Functions

The second novel feature integrated into the proposed fuzzy snake is the fuzzy representation of energy functions. Using linguistic variables, this approach provides the active contour with an intuitive man-machine-interface, allowing uncertain knowledge to be exploited.

4.1 Fuzzification

With the fuzzy snake, the components of the external energy function (equation (2)) are separately represented by linguistic variables and fuzzy sets.

With E_{image} the usual calculation of image evidences for edges, regions or the motion of an object [7] is retained, but they become linguistic variables.

The example in Fig. 3 illustrates how an evidence function can be extended by an intuitive interface using elements of fuzzy logic. For a normalise gradient function a linguistic variable *edge* is created. Fuzzy sets with linguistic values, for instance falling very strong, ..., rising very strong are defined.[1] In an analogous manner, the definition of linguistic variables and fuzzy sets for region and motion-based evidences are similarly given.

To illustrate how a constraint energy function is integrated into the fuzzy snake, the fuzzification of a curvature assessing function is shown in Fig. 4. The geometrical constraint becomes a linguistic variable *curvature* which is calculated using an empirically determined mapping function between linguistic terms and a local angle of the boundary, which represents the actual measure for the constraint. The corresponding polynomials of the membership functions may be approximated by piecewise linear functions to reduce computational complexity.

Many other alternative constraints may also be included, for example the orientation of a contour segment (by using the angle of the major axis of a segment) or the relative position of different contour segments.

The entire membership degrees $\mu(edge)$, $\mu(region)$ and $\mu(motion)$ for a given candidate $c_{n,m}$ can thus be regarded as its fuzzy image energy, and membership degrees for constraints such as $\mu(curvature)$ correspondingly as its fuzzy constraint energy.
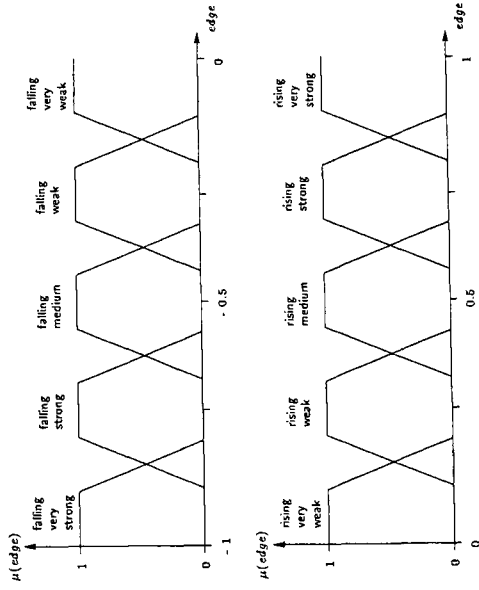

Figure 3: Fuzzy representation of edge-based image evidences.
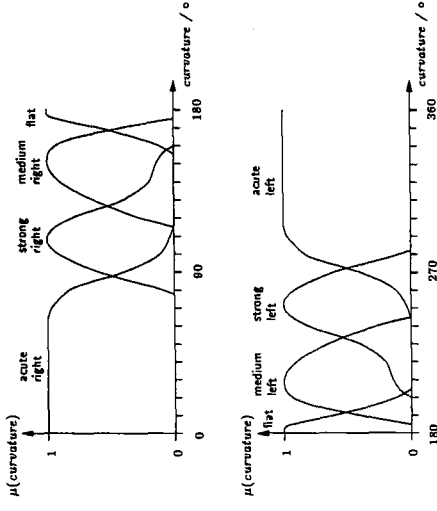
Figure 4: Fuzzy representation of curvature constraint.

4.2 Linguistic rules

Algebraic equations commonly used for specifying E_{image} and E_{con} can now be replaced by a fuzzy inference process using formal linguistic rules instead of trial-and-error weights.

As previously described, each energy function for a boundary segment exhibiting constant properties consists of a rule base, which inputs a number of different evidences or features from the image, as well as constraints on for example, the geometry of an object. The output from the inference process using this rule base, is a *quality* measure (Fig. 5), describing the compliance with the rule base for each vertex. Normally rules will have the conclusion *quality* very good to describe a known desired result, while negative linguistic values for *quality* may be used for those properties which a contour segment must *not* exhibit.

Considering the edge and curvature properties of the example shown in Fig. 2, the contour can now be described by the four rule bases a, b, c and d in Table 2.

Table 2: Rule bases for the example in Fig. 2.

a:
IF *edge* falling very weak
AND *curvature* medium right
THEN *quality* very good

b:
IF *edge* falling weak
AND (*curvature* flat left
OR *curvature* flat right)
THEN *quality* very good

c:
IF *edge* falling medium
AND *curvature* strong right
THEN *quality* very good

d:
IF *edge* falling medium
AND (*curvature* flat left
OR *curvature* flat right)
THEN *quality* very good

The defuzzification procedure converts the result of the inference process which is a fuzzy quality measure, into a crisp value $E^*_{ext}(c_{n,m})$. Here the well-referenced centre-of-gravity method [9,10,11] is used for defuzzification.

[1]Note for simplicity, that the adjectives are regarded as a part of the primary term of the linguistic value, rather than as a linguistic hedge with an associated operator in the sense of [8].
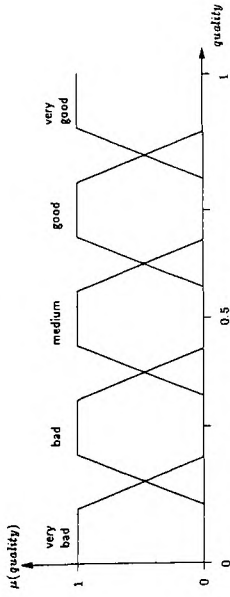
Figure 5: Output variable of the fuzzy system.

5 Example Application of Fuzzy Snakes

One particular application area which has proved especially beneficial for exploiting the concept of fuzzy snakes in preference to more traditional active contouring techniques is the field of medical imaging.

A specific example is the segmentation of wrist bones in magnetic-resonance image (MRI) sequences. Here the contours of a number of bones are to be segmented precisely in order to measure the objects' 2D motion, while the patient's wrist bends. All objects are known, though their characteristic shape varies individually.
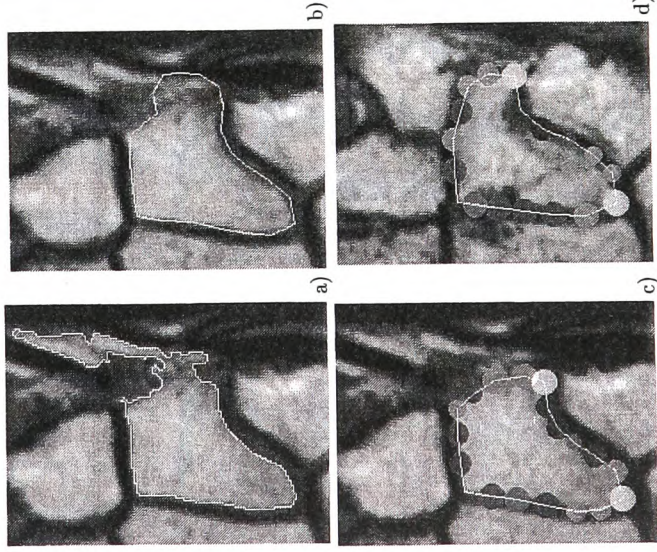

Figure 6: Improved segmentation through fuzzy snakes. a) histogram-based approach presented in [12], b) traditional active contour with global shape constraint, c) and d) segmentation in images of two different patients using fuzzy snakes.

Image segmentation methods based primarily on low-level image features obtain good results in many cases [12], though crucially where neighbouring anatomic structures are too similar with respect to their features, then segmentation fails. In Fig. 6a) for example, a tissue region not belonging to the bone was segmented in the upper right area of the image. These errors can only be corrected through integration of *a priori* knowledge concerning the shape of the object.

Traditional active contours use only global constraints, which improve the segmentation though this is insufficient in certain cases. In Fig. 6b) the global shape constraint is able to create a smooth contour resulting in the exclusion of most of the false tissue. The right-hand area of the result however, is still incorrect as the active contour is attracted to strong image features that could not be overridden by the moderate global shape constraint. A stronger influence of the constraint in this critical area is necessary.

Fuzzy snakes allow for a more detailed modelling of the object, resulting in a correct segmentation. The results in Fig. 6c) were obtained with the description (using relative fuzzy lengths) $(0.2d)(0.1b)(0.05c)(0.1e)(0.15f)(0.05c)(0.05f)(0.3a)$ beginning at the upper left vertex and following the contour clockwise. The characters are labels for the rule bases given in Table 3.

Table 3: Rule bases modelling the boundary segments of a wrist bone (Os Hamatum) in MRI sequences as shown in Fig. 6.

a:
IF *edge rising medium*
AND (*curvature flat left*
OR *curvature flat right*)
THEN *quality very good*

b:
IF *edge rising medium*
AND *curvature medium right*
THEN *quality very good*

c:
IF *edge rising medium*
AND *curvature strong right*
THEN *quality very good*

d:
IF *edge rising strong*
AND (*curvature flat left*
OR *curvature flat right*)
THEN *quality very good*

e:
IF *edge rising strong*
AND *curvature medium left*
THEN *quality very good*

f:
IF *edge rising strong*
AND *curvature medium right*
THEN *quality very good*

The fuzzy snake is able to handle a certain variability in the object's contour, which allows for the segmentation of a bone over a patient's MRI sequence. Furthermore, since the inter-individual variations tend to be moderate, the bones of other patients can be segmented successfully using the same contour description, or prototype as shown in Fig. 6d).

6 Conclusions

This paper has introduced the concept and presented the theoretical basis of the fuzzy snake as a new form of active contour, broadening the areas of application of this well known contour identification approach. It has been shown that the principal advantage of this new method in particular, is the ability to exploit uncertain *a priori* knowledge, such as a verbal description of contour properties. The proposed new fuzzy energy functions present a level of abstraction which is higher and therefore closer to the human expert than that of algebraic energy functions. Parameters and weights with sometimes

obscure meanings are replaced by the more intuitive linguistic interface provided by fuzzy logic expressions.

By allowing a more detailed object description it has been proved that the fuzzy snake approach can improve boundary detection in complex images of poor quality while concomitantly reducing computational complexity. An appropriate real-world example has been presented to validate the performance in being able to identify structures in the processing of MRI based imaging sequences.

References

[1] M. Kass, A. Witkin, and D. Terzopoulos, Snakes: Active contour models, International Conference on Computer Vision, London, 1987, pp. 259–268

[2] A. A. Amini, S. Tehrani, and T. E. Weymouth, Using dynamic programming for minimizing the energy of active contours in the presence of hard constraints, 2nd International Conference on Computer Vision, 1988, pp. 95–99

[3] B. Olstad, Active contours with grammatical descriptions, 6th International Conference on Image Analysis and Processing, Como, Italy, September 1991

[4] B. Olstad and A. H. Torp, Encoding of a priori Information in Active Contour Models, *Pattern Analysis and Machine Intelligence*, **18** (9), 1996, pp. 863–872

[5] F. Höwing, L. S. Dooley, and D. Wermser, Linguistic contour modelling through Fuzzy Snakes, Computational Intelligence for Modelling Control and Automation, M. Mohammadian (ed.), IOS Press, Amsterdam, 1999, pp. 384-389

[6] J. C. Bezdek and S. K. Pal (eds.), Fuzzy models for pattern recognition, IEEE Press, New York, 1992

[7] F. Höwing, L. S. Dooley, and D. Wermser, Tracking of non-rigid articulatory organs in X-ray image sequences, *Computerized Medical Imaging and Graphics*, **23** (2), 1999, pp. 19–27

[8] L. A. Zadeh, Fuzzy Sets, *Inform. Control*, **8**, 1965, pp. 338–353

[9] H. Bandemer, Einführung in Fuzzy-Methoden, Akademie-Verlag, Berlin, 1993

[10] B. Kosko, Neural Networks and Fuzzy Systems, Prentice-Hall, Englewood Cliffs, N.J., 1992

[11] T. Tilli, Mustererkennung mit Fuzzy-Logik, Franzis-Verlag, München, 1993

[12] F. Höwing, H. Bülow, D. Wermser, L. S. Dooley, and W. Thoma, Automatic Motion Analysis of Bones from MR Sequences, International Conference on Image Processing and its Applications, Manchester, UK, July 1999, pp. 397 – 401

Fuzzy active contour model

F.Höwing, L.S.Dooley and D.Wermser

Abstract: A new method for representing and tracking of object boundaries is presented, which allows for the integration of uncertain *a priori* knowledge into an active contour model. The novel concept of fuzzy snakes is developed to allow for an intuitive specification of the properties of an object's boundary. This is achieved by introducing fuzzy energy functions and establishing a linguistic rule base, which describes each of the fuzzy snake's segments. Furthermore the approximate length of each contour segment may be specified to both improve the segmentation process and to reduce computational complexity. Experimental results demonstrate the validity of the theoretical properties of the fuzzy snake approach, and examples have been included illustrating the application of the technique to complex scenes, such as medical imaging sequences.

1 Introduction

Active contours [1], or 'snakes', are a well known method for matching an object's contour model to features in an image. The approach, which uses a polygonal object representation (Fig. 1), is distinguished by its intrinsic ability to handle variations in the boundary that is to be detected. It is therefore capable of identifying and tracking deformable objects in image sequences.

For each image, the algorithm requires an initial polygon $P = (p_0, p_1, \ldots, p_{N-1})$ consisting of N vertices $p_i = (x_i, y_i)$, where x_i and y_i are the spatial co-ordinates of p_i. The detected boundary is represented by the polygon $Q = (q_0, q_1, \ldots, q_{N-1})$, with $q_i = (x_i, y_i)$. Each q_i is selected from a set of candidates $C_i = (c_{i,0}, c_{i,1}, \ldots, c_{i,M-1})$. In many applications the candidates $c_{i,j} = (x_j, y_j)$ are uniformly sampled along a search line normal to the initial polygon and intersecting p_i. Tracking of the contour is achieved by processing a sequence frame by frame and taking the resulting Q_t as the initial estimated contour P_{t+1} for the next frame.

From this polygonal representation, it is possible to formulate an appropriate energy function E_{snake} for the object, which can subsequently be minimised to obtain the desired contour Q. This minimisation is achieved by selecting an optimal set of vertices from the candidates which have been sampled in a region R, around an initial contour. Normally E_{snake} is decomposed into two components:

$$E_{snake}(Q) = \sum_{n=0}^{N-1} w_{int}E_{int}(q_n) + w_{ext}E_{ext}(q_n) \qquad (1)$$

E_{int}, the internal energy, represents general contour properties such as stiffness, which are usually required to produce a smooth shape. The external energy E_{ext} is composed of E_{image}, which guides the contour towards particular features in the image, and E_{con}, which allows for the integration of additional constraints (cf. eqn. 2). Weighting parameters w_{int}, w_{ext}, w_{image} and w_{con} control the relative influence of the energy components and are generally determined by a process of trial-and-error,

$$E_{ext}(q_n) = w_{image}E_{image}(q_n) + w_{con}E_{con}(q_n) \qquad (2)$$

The original active contour algorithm presented in [1] had some inherent computational problems in evaluating the energy function, which were subsequently solved by Amini *et al.* [2]. The energy minimisation of the snake was performed by a discrete dynamic programming algorithm which allowed for the integration of hard constraints, such as a minimum distance between the snake's vertices.

A simplified version of the dynamic programming algorithm is given in Table 1, where N is the number of vertices and M the number of candidates for each vertex. $S(n, m)$ represents the minimal energy level possible for the vertices $0, \ldots, n$ if the nth vertex is the candidate

Table 1: Dynamic programming algorithm for optimisation of a global energy function

1.   **for** $n = 1 \ldots N - 1$
2.   **for** $m = 0 \ldots M - 1$
3. $S(n, m) = \min_k [w_{int}E_{int}(c_{n,m}) + w_{ext}E_{ext}(c_{n,m}) + S(n-1, k)]$
4. $T(n, m) = k^{min}$
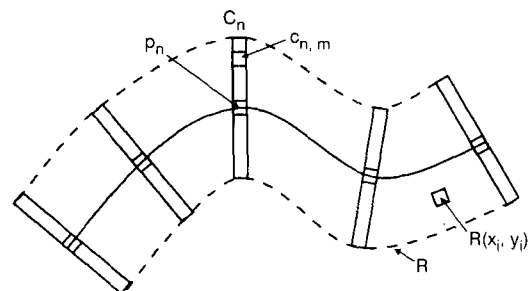
Fig. 1 *Polygonal contour representation*

$c_{n,m}$, k^{min} is the $k \in [0, M-1]$ that minimised the expression in line 3. It points to the optimal predecessor of $c_{n,m}$ and is stored in $T(n, m)$. After all vertices have been processed, these pointers are traced back to obtain the new boundary Q. Olstad [3] successfully applied this algorithm to the detection of the left ventricle in ultrasonic images.

An alternative optimisation strategy allowing for hard constraints is the greedy algorithm proposed by Williams and Shah in [4], while an extensive survey of other optimisation strategies which also have been applied to active contours is given in [5]. Amongst some of the principal items raised are issues concerning the large number of iterations and convergence problems. In both these respects the approach proposed in this paper requires only a single iteration to find the desired object boundaries, provided they were within the search lines.

Many of the approaches adopted are concerned with implicit knowledge representation. Examples include statistical models of shape in [6, 7], where active contours are trained from a set of examples, and the most popular and effective approach of active shapes in [8] is based on a point distribution model.

The fuzzy active contour presented in this paper is based on an explicit representation of an object's properties, including but not limited to shape. It is appropriate in situations where a representative training set is either not available or where an explicit and sometimes verbal knowledge representation is more appropriate in the context of the application. For instance, in medical domain applications selected non-computer vision experts can easily describe both shape *and* appearance in linguistic terms.

2 Active contours with multiple segments

In traditional active contours a fundamental limitation remains in terms of representing *a priori* knowledge concerning more complex objects which are to be detected. This is that all vertices of the active contour are characterised by the same local energy function, resulting in a single, global description of the object. To overcome this problem, Olstad [9, 10] introduced a grammatical description of the snake's energy function. Fig. 2 shows an example of a contour, which can be described as a sequence of four different external energy functions E_{ext}^{a}, E_{ext}^{b}, E_{ext}^{c} and E_{ext}^{d}, represented by the respective terminals a, b, c and d. The grammatical expression describing such a segmented boundary would be $\alpha = a*b*c*d*$, where * is
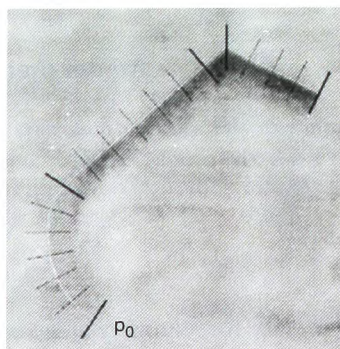

Fig. 2 *Example object*

This synthetic boundary demonstrates a multisegment object contour. The boundary segments have a different shape (local curvature) as well as a different appearance in the image (edge contrast), while all vertices within a segment share similar properties. The contour's vertices p_n are visualised by lines perpendicular to the contour. Thick lines illustrate the segment boundaries
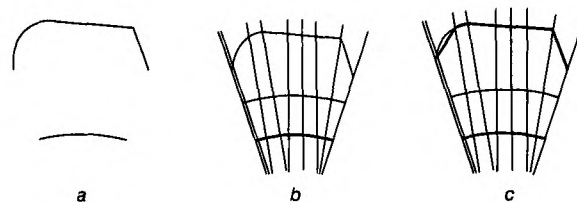

Fig. 3 *Two objects with different curvature properties*

Lower constant contour can be detected by a traditional active contour (b), while upper multisegment contour requires a better model. Figs. b and c also contain initial contour hypothesis *P* (upper arc) as well as search lines C_i (straight lines)

the closure operator which allows parts of a pattern to be arbitrarily repeated. A pattern-matching algorithm is incorporated within the active contour's energy minimisation process which constrains the resulting contours to only those that comply with the grammatical expression.

To illustrate the benefits of the grammatical model consider only the *local shape* (curvature) of the example object in Fig. 2 resulting in the upper object in Fig. 3a. The traditional active contour is only capable of representing *constant* local properties. Its similarity constraint favours the constant contour (Fig. 3b). With the grammatical model the more complex contour can be represented and hence detected (Fig. 3c).

3 Fuzzy contour model

The grammatical approach described in the preceding section may considerably increase the number of energy functions, their parameters and weights. Often these numerical parameters are obscure and their refinement time consuming. This is especially true in situations where the *a priori* knowledge is imprecise or uncertain, or where it is explicitly available in the form of verbal expressions. It then becomes much more desirable to use a more intuitive contour description.

This section introduces the general concept of a fuzzy-logic based active contour model [11, 12]. In addition to a representation where the length of each contour segment can be specified it permits a linguistic description of various properties of an object's boundary at a high level of abstraction.

Analogous to the image energy E_{image} of an active contour, the fuzzy snake is capable of representing the appearance of an object in the image. Further properties of the object's boundary segments, such as shape, may be represented in a similar way to the function E_{con} in an active contour. The model, however, uses linguistic variables and linguistic values of fuzzy sets instead of numerical parameters.

To illustrate this approach consider the local shape of the example object in Fig. 2 again. The boundary may be decomposed into four segments. With Olstad's model, a grammatical expression to describe this contour could be $\alpha = a*b*c*d*$, where a represents an energy function favouring local angles of $160°$, while b and c favour angles of $180°$ and $110°$, respectively.

A more intuitive description would be 'a medium length arc, bending right, followed by a medium straight line, a right bending corner and a short straight line'. Such a verbal expression can be formalised through the proposed fuzzy contour description D, which is a concatenation of segment descriptions d_z, $z \in [0, Z]$, with Z being the number of boundary segments. Each d_z is decomposed into a fuzzy segment length l_z and a segment property z_z.

Table 2: Example of a linguistic fuzzy contour description in terms of each segment's length and property

D = *medium* right arc, *medium* straight line
very short right corner, *short* straight line

Table 3: Fuzzy contour description using fuzzy lengths and shortcut property labels

D = (6a)(6b)(1c)(3d)

Table 4: Property labels defined by linguistic rules

Segment property label	Shortcut label	Fuzzy rule base
right arc	a	IF *curvature* medium right THEN *quality* very good
straight line	b	IF *curvature* flat left OR *curvature* flat right THEN *quality* very good
right corner	c	IF *curvature* strong right THEN *quality* very good

Table 2 demonstrates the use of linguistic values for l_z (*medium, short* and *very short*). To integrate them into an algorithm, they are translated into a fuzzy number of snake vertices (cf. Table 3). The details of this mapping are discussed in Section 4. The property labels for z_z used in Table 2 (right arc, straight line and right corner) represent linguistic fuzzy rules to describe all the features a contour may exhibit. For convenience short-cut labels may be used (cf. Tables 3 and 4). To realise this, a linguistic variable is created for each feature. To describe a local shape, for example, the local angle at each vertex is measured and mapped to a linguistic variable *curvature*. Fuzzy sets are created and linguistic values assigned to characterise the curvature as, for example, acute right or flat. Taken together with an output variable *quality*, it is then possible to describe each segment's curvature by a fuzzy rule base. An example is given in Table 4.

The calculation of the active contour's external energy is performed by a fuzzy inference [13], the defuzzified [14–16] output of which is a crisp quality measure for each contour candidate. The details of this approach are presented in Section 5.
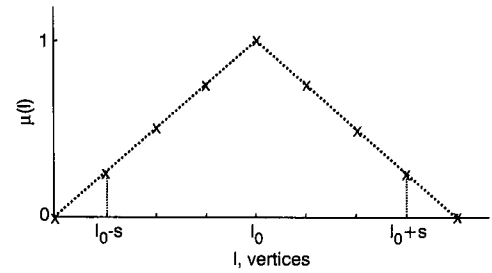
4 Fuzzy segment length

The fuzzy snake approach permits the specification of the number of subsequent vertices that share a common energy function, a parameter constraint which dramatically reduces the search space.

4.1 Notation

Since snake segment lengths are not known precisely, a new notation is introduced, where the length is expressed as a fuzzy number (cf. Table 3). Fig. 4 illustrates such a (discrete) fuzzy number [17] l, with mean value l_0 and spread s, defined by $\mu(l) > 0 \; \forall l \in \{l_0 - s, \ldots, l_0 + s\}$.

In the first segment of the contour in Fig. 2, for a fuzzy length of $l_0 = 6$ and a spread of $s = 2$, the segment speci-

Fig. 4 *Segment length as a fuzzy number*

fication (6a) denotes a segment consisting of between four and eight vertices, sharing a common property defined by an energy function E^a_{ext}. Using the analogy of a grammatical description, the fuzzy snake can now be specified as an expression, using operators such as concatenation, logical AND and OR, together with a fuzzy length.

The absolute number of vertices depends on the sampling distance and is not a direct measure for the *length* of a segment. While the fuzzy snake algorithm uses the absolute length, the user-level contour description must allow for a length measure relative to the overall length of the contour. With the segment length expressed as a percentage, the above example can be written as (*0.4*a) (*0.4*b) (*0.05*c) (*0.15*d). These relative lengths can now be mapped to linguistic labels as shown in Table 5 to realise a contour description as proposed in Table 2.

4.2 Extended optimisation algorithm

An extension to the dynamic programming algorithm outlined in Section 1 is necessary to account for the *variable* length of snake segments. Multiple external energy functions imply that many different calculations have to be performed for each vertex. Since the segment length is now variable, several energy functions may be calculated simultaneously for vertices either on or near a segment boundary. To illustrate this, an example is provided in Table 6.

When back-tracking takes place, only one energy function for each vertex succeeds. The selected function thereby determines the final *state* of that vertex.

The actual implementation of the fuzzy segment length is achieved by an extension to the expression in line 3 of Table 1. The new expression is rewritten in line 3 of Table 7. The

Table 5: Mapping of linguistic labels to fuzzy numbers: example for expression in Table 2 with N = 16

Linguistic label	Relative length	Absolute fuzzy length
very short	≈ 5%	1
short	≈ 15%	3
medium	≈ 40%	6

Table 6: Energy functions which have to be considered in the Fig. 2 example

Vertex	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Functions	a	a	a	a	a	a	a	a	a	a								
			b	b	b	b	b	b	b	b	b	b						
					c	c	c	c	c	c	c							
							d	d	d	d	d	d	d	d				

Table 7: Extended dynamic programming algorithm

1.	**for** $n = 1 \ldots N - 1$
2.	**for** $m = 0 \ldots M - 1$
3.	$S_z(n, m) = \min_k [w_{int} E_{int}(c_{n,m}) + w_{ext} E_{ext}^z(c_{n,m}) + S_z(n - 1, k) + 1 - \mu(l_z(n - 1, k))]$
4.	$l_z(n, m) = l_z(n - 1, k^{min}) + 1$
5.	$T_z(n, m) = k^{min}$

length is regarded as an additional constraint, where $l_z(n - 1, k)$ is analogous to $S_z(n - 1, k)$ and denotes the number of preceding vertices which would fall into state z if candidate $c_{n,m}$ was selected. Consequently, candidates which assist in constructing a chain of the specified length are favoured. Line 4 of Table 7 updates the length information. T_z in line 5 points to that predecessor of $c_{n,m}$ which would be optimal if the final state of vertex n was z.

An additional advantage occurring from this approach is that computational complexity is reduced considerably. In the above example an arbitrary segment length would generate $(N - 1)^{(Z - 1)} = 15^3 = 3375$ possible states of the contour, where Z is the number of segments. Exploiting the *a priori* knowledge concerning the approximate segment length reduces this number by approximately 98.5% to 52. It must be stated, however, that the actual computational complexity is very much application-specific and depends on l_z, s and the complexities $O(E_{ext}^z)$, where $Z \in \{a, b, c, \ldots\}$.
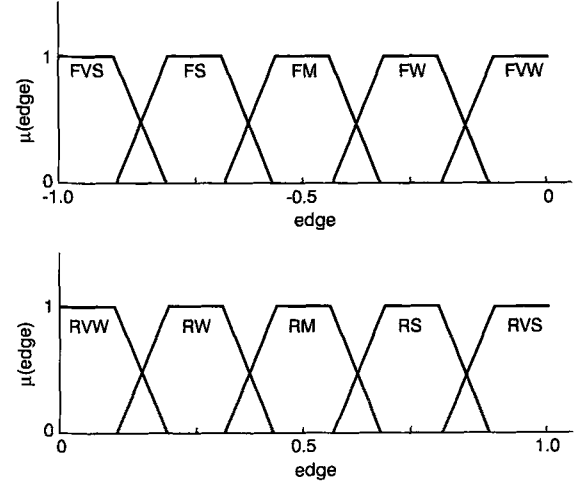
5 Fuzzy energy functions

The second novel feature integrated into the proposed fuzzy snake is the fuzzy representation of energy functions. Using linguistic variables, this approach provides the active contour with an intuitive man-machine interface, allowing uncertain knowledge to be exploited.

5.1 Fuzzification

The Appendix defines an example set of algebraic energy functions typically used by a traditional active contour. With the fuzzy snake, the components of the external energy function (eqns. 4 and 8) are represented separately by linguistic variables and fuzzy sets. For E_{image} the calculation of image evidences (eqns. 5–7) is retained, but they become linguistic variables.

The example in Fig. 5 illustrates how an evidence function can be extended by an intuitive interface using elements of fuzzy logic. For e_{edge}, a linguistic variable *edge* is created. Fuzzy sets with linguistic values, for instance falling very strong,..., rising very strong, are defined to cover the value range of eqn. 5 [Note 1]. In this example, the values of the term $V_\perp(C_n, c_{n,m})$ are mapped to adjectives which cover the range from very weak to very strong. The factor δ, which specifies the edge direction, is accounted for by the adjectives rising and falling, respectively.

In an analogous manner, the definition of linguistic variables and fuzzy sets for the region and motion-based evidences (eqns. 6 and 7, respectively) are similarly given. To illustrate how a constraint energy function is integrated into the fuzzy snake, the fuzzification of eqn. 8 is shown in Fig. 6. Although eqn. 8 already allows for the integration of inexact knowledge, it affords a number of parameters

Note 1: For simplicity, that the adjectives are regarded as a part of the primary term of the linguistic value rather than as a linguistic hedge with an associated operator in the sense of [18].

Fig. 5 *Fuzzy representation of edge-based image evidences*

F, falling; R, rising; VS, very strong; VW, very weak; S, strong; M, medium, W, weak
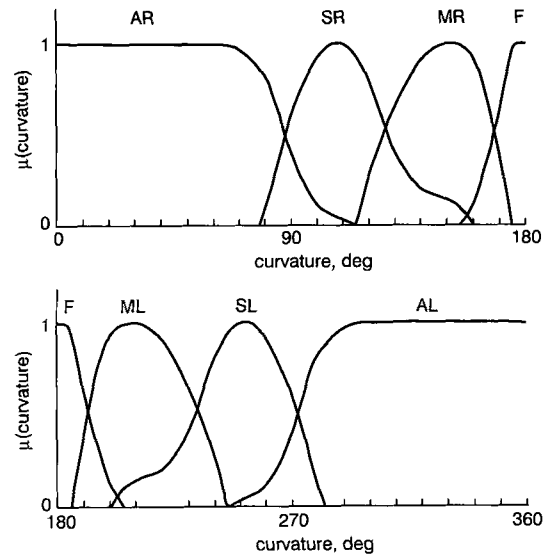
Fig. 6 *Fuzzy representation of curvature constraint*

AR, acute right; SR, strong right; MR, medium right; F, flat; ML, medium left; SL, strong left; AL, acute left

and hence does not provide an intuitive interface to either the shape constraint or other constraints that may similarly be applied. Through fuzzification the geometrical constraint becomes a linguistic variable *curvature* which is calculated using eqn. 9 and represents the actual measure for the constraint.

The weighting function v in eqn. 10, however, is now replaced by a number of fuzzy sets. The membership functions of the fuzzy sets acute to flat have been determined empirically to relate to the human perception of the different degrees of curvature. Many other alternative constraints may also be included, for example the

orientation of a contour segment (by using the angle of the major axis of a segment) or the relative position of different contour segments.

The entire membership degrees $\mu(edge)$, $\mu(region)$ and $\mu(motion)$ for a given candidate $c_{n,m}$ can thus be regarded as its fuzzy image energy [Note 2], and membership degrees for constraints such as $\mu(curvature)$ correspondingly as its fuzzy constraint energy.

5.2 Fuzziness of boundary features

The fuzzy snake model constrains the possible shapes the resulting contour might have, while concomitantly allowing for a certain deviation from an optimal prototype of the object that is to be detected. The variable segment length is one element which affords this flexibility. Another is the fuzziness in the description of desired boundary features. To illustrate how a fuzzy snake is capable of favouring desired features using the fuzzification described above, the following example visualises the membership degrees of the fuzzy sets over the linguistic variable *curvature*.

For each vertex of the example contour in Fig. 2, all membership degrees for the fuzzy sets in Fig. 6 are calculated. The highest and second highest value (if any), for each vertex are recorded in the graph in Fig. 7. The graph shows that the analysed contour is very similar to the description given in Tables 2 and 4, which demand curvatures of medium right, flat left OR flat right, strong right, flat left OR flat right. These relations become more evident when a different visualisation is applied as demonstrated in Fig. 8. Here the fuzzy sets are colour coded and superimposed as circles over the contour image. The circle diameter is proportional to the
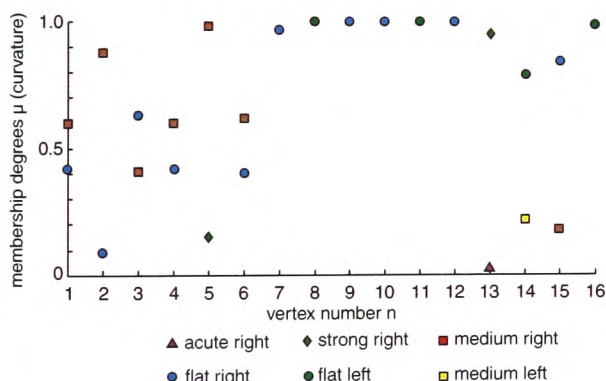

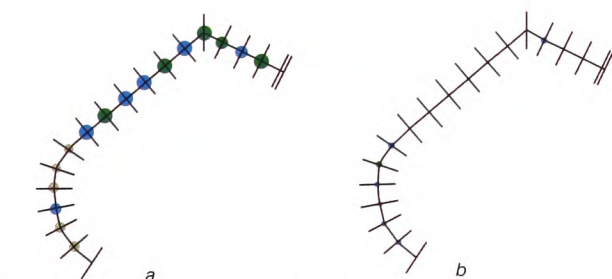**Fig. 7** *Significant membership degrees for each vertex of a contour*

Fig. 8 *Visualisation of membership degrees for boundary features*

Fuzzy sets are colour coded according to Fig. 7
a Maximum membership degrees of each vertex
b Second highest values

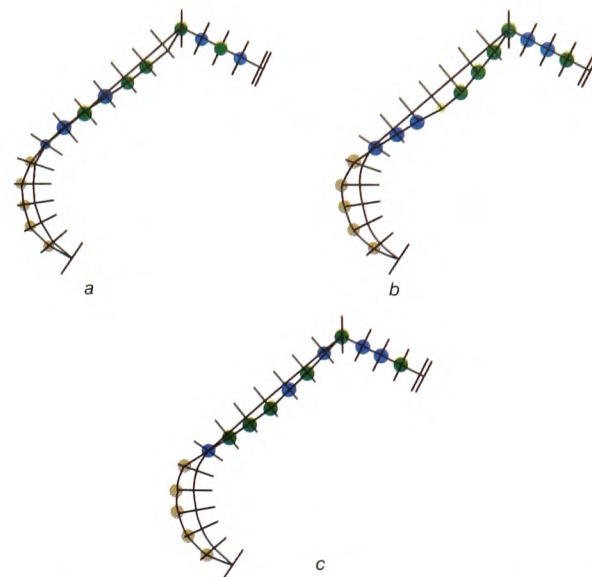Note 2: It should be recalled that high membership values relate to low energy values (see definition of eqn. 4.)

Fig. 9 *Demonstration of shape feature's fuzziness through a visualisation of mean maximum membership degrees for curvature of different nonrigid deformations of a prototype*

mean degree of membership within a consecutive sequence of equal fuzzy sets.

As both the representation and detection of non-rigid objects are of particular interest, a sequence of deformations of the example contour is shown in Fig. 9. This sequence demonstrates to what degree the fuzziness of a contour description favours certain deviations from the prototype while placing less emphasis on others. In Fig. 9a–c a local distortion was introduced in the longer straight segment, bending the contour to the left. With the deformations in Figs. 9a and b the local curvature at the distorted vertices results in a high value for $\mu_{\text{medium left}}(curvature)$, while the desired $\mu_{\text{flat left}}(curvature)$ is very small or zero (not shown). This means that the overall energy E_{snake} is significantly smaller for Figs. 9a and b than it is for Fig. 9c. In other words the deformation shown in Fig. 9c is more similar to the prototype than that in Fig. 9a or b.

5.3 Linguistic rules

As previously described, each energy function for a boundary segment exhibiting constant properties consists of a rule base, which inputs a number of different evidences or features from the image, as well as constraints on, for example, the geometry of an object. The output from the inference process using this rule base is a *quality* measure (Fig. 10), describing the compliance with the rule base for each vertex.

Normally rules will have the conclusion *quality* very good to describe a known desired result, while negative
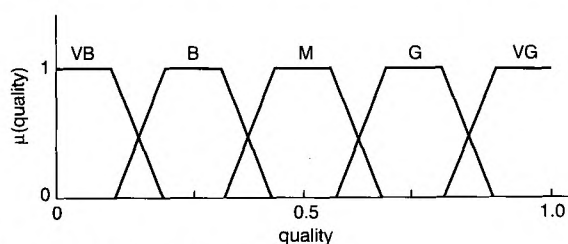

Fig. 10 *Output variable of fuzzy system*
VB, very bad; B, bad; M, medium; G, good; VG, very good

327

Table 8: Rule bases for example in Fig. 4

a:	c:
IF *edge* falling very weak	IF *edge* falling medium
AND *curvature* medium right	AND *curvature* strong right
THEN *quality* very good	THEN *quality* very good
b:	d:
IF *edge* falling weak	IF *edge* falling medium
AND (*curvature* flat left	AND (*curvature* flat left
OR *curvature* flat right)	OR *curvature* flat right)
THEN *quality* very good	THEN *quality* very good

linguistic values for *quality* may be used for those properties which a contour segment must *not* exhibit.

Considering the edge and curvature properties of the example shown in Fig. 2, the contour can now be described by the four rule bases a, b, c and d in Table 8.

The defuzzification procedure converts the result of the inference process, which is a fuzzy quality measure, into a crisp value $E^z_{ext}(c_{n,m})$. Here the well referenced centre-of-gravity method [14–16] is used for defuzzification.

6 Application examples of fuzzy snakes

An application area which has proved beneficial for exploiting the concept of fuzzy snakes in preference to more traditional active contouring techniques is the field of medical imaging. In this paper two specific applications are presented. First, the segmentation of wrist bones from MRI image sequences [19] and secondly the tracking of the tongue from moving X-ray sequences of the oral cavity [20, 21].

6.1 Segmentation of wrist bones in magnetic-resonance image (MRI) sequences

Here the contours of a number of bones are to be segmented precisely to measure the objects' 2-D motion while the patient's wrist bends. All objects are known, though their characteristic shape varies individually.

Image segmentation methods based primarily on low-level image features obtain good results in many cases [19], though crucially where neighbouring anatomic structures are too similar with respect to their features then segmentation fails. In Fig. 11a, for example, a tissue region not belonging to the bone was segmented in the upper right area of the image. These errors can only be corrected through integration of *a priori* knowledge concerning the shape of the object.

Traditional active contours use only global constraints, which improve the segmentation, but this is not sufficient in certain cases. In Fig. 11b the global shape constraint is able to create a smooth contour, resulting in the exclusion of most of the false tissue. The right-hand area of the result, however, is still incorrect as the active contour is attracted to strong image features that could not be overridden by the moderate global shape constraint. A stronger influence of the constraint in this critical area is necessary. Note that, had the initial contour been closer to the bone, the result could be improved. However, such an initial contour would then be too specialised to be applied to more than one patient – an issue revisited later in this section.

Fuzzy snakes allow for a more detailed modelling of the object, resulting in a correct segmentation. The results in
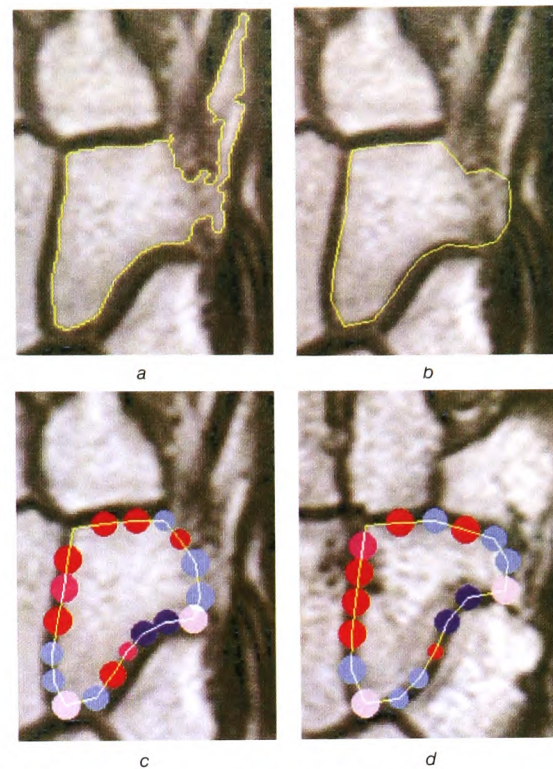
Fig. 11 *Improved segmentation through fuzzy snakes*
a Histogram-based approach presented in [19]
b Traditional active contour with global shape constraint
c, d Segmentation in images of two different patients using fuzzy snakes

Fig. 11c were obtained with the description (using relative fuzzy lengths) $(0.2d)$ $(0.1b)$ $(0.05c)$ $(0.1e)$ $(0.15f)$ $(0.05c)$ $(0.05f)$ $(0.3a)$ beginning at the upper left vertex and following the contour clockwise. The characters are shortcut labels for the rule bases given in Table 9. The description was derived from a linguistic description of the object provided by orthopedic practitioners.

The fuzzy snake is able to handle a certain variability in the object's contour, which allows for the segmentation of a bone over a patient's MRI sequence. Furthermore, the inter-individual variations are moderate, so the bones of other patients can be segmented successfully using the same contour description, or prototype as shown in Fig. 11d.

Table 9: Rule bases modelling boundary segments of a wrist bone (Os Hamatum) in MRI sequences as shown in Fig. 11

a:	d:
IF *edge* rising medium	IF *edge* rising strong
AND (*curvature* flat left	AND (*curvature* flat left
OR *curvature* flat right)	OR *curvature* flat right)
THEN *quality* very good	THEN *quality* very good
b:	e:
IF *edge* rising medium	IF *edge* rising strong
AND *curvature* medium right	AND *curvature* medium left
THEN *quality* very good	THEN *quality* very good
c:	f:
IF *edge* rising medium	IF *edge* rising strong
AND *curvature* strong right	AND *curvature* medium right
THEN *quality* very good	THEN *quality* very good

6.2 Segmentation and tracking of tongue in X-ray image sequences

In the second example application, articulatory organs have to be measured to investigate the complex dynamics of speech production. Of particular interest is the robust boundary detection of the tongue in sagittal X-ray imaging sequences, which poses particular problems to image processing.

Medical X-ray sequences are characterised by the presence of transparently superimposing structures and varying textural appearances of organs and noise. Under these conditions local low-level based image processing operators are not able to segment an object such as the tongue. It was demonstrated by Meyer [22] that at least a minimum of local shape information is required to obtain first results in single X-ray images (cf. Fig. 12*a*).

In [21], satisfactory results for many positions of the tongue were presented (cf. Fig. 12*b*), emphasising the need for combined image features as well as a shape constraint as outlined in the Appendix. To detect several different shapes of the tongue it was necessary to allow for a relatively high tolerance in the shape constraint, but this however reduced the robustness of the active contour against distortions. Furthermore, the tip of the tongue was not detected as its locally higher curvature could not be modelled through the global shape constraint.

These problems were overcome using the more detailed contour description afforded by the fuzzy snake. The description in Table 10, together with the rule bases in Table 11, allowed for a modelling of the segments of the tongue for a characteristic position. Locally different curvatures were accounted for, increasing the robustness of the segmentation as well as enabling the detection of the

Table 10: Contour description of one characteristic position of tongue

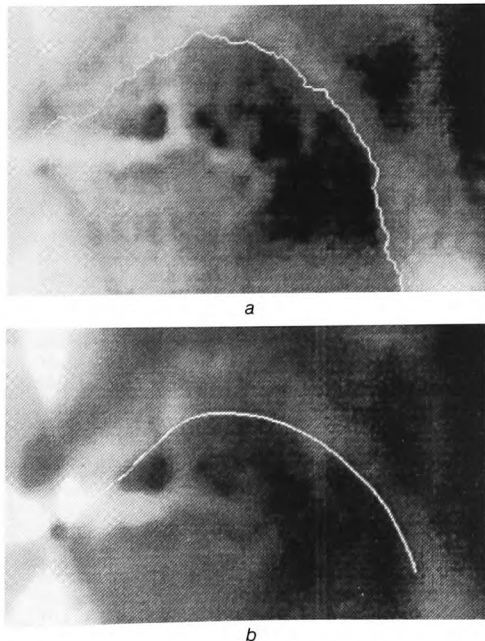very short tip, short flat front, medium bent middle, medium bent back

Fig. 12 *Segmentation of tongue through local contour tracing and traditional active contours*

a Local contour trace
b Traditional active contours

Table 11: Some rule bases for modelling boundary segments of tongue

| tip:	flat front:
IF *edge* falling very weak	IF *edge* falling medium
AND *motion* medium	AND *motion* strong
AND *curvature* strong right	AND *curvature* flat right
THEN *quality* very good	THEN *quality* very good
IF *edge* falling medium	IF *region* negative strong
OR *edge* falling strong	OR *region* negative very
THEN *quality* bad	strong
	THEN *quality* bad
bent middle:	bent back:
IF *edge* falling weak	IF *edge* falling medium
AND *motion* strong	AND *motion* medium
AND (*curvature* flat right	AND (*curvature* flat right
OR *curvature* medium right)	OR *curvature* medium right)
THEN *quality* very good	THEN *quality* very good
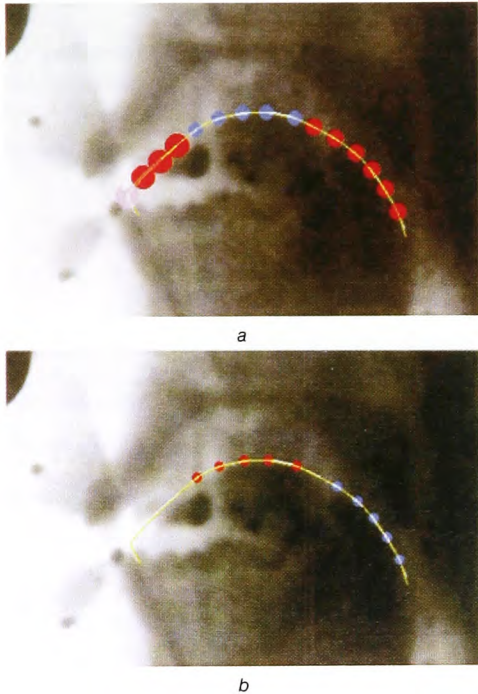
Fig. 13 *Fuzzy snake detection of tongue with superimposed mean membership degrees for curvature fuzzy sets*

a Highest values
b Second highest values

tip of the tongue (cf. Fig. 13). Structures that were able to distort the snake were explicitly modelled, increasing the correctness of the segmentation. The second rule in the rule base tip, for instance, accounts for the lower jaw, while the IF-*region* rule in the rule base flat front reduces the influence of a tooth filling located in the upper jaw.

7 Conclusions

This paper has introduced the concept and presented the theoretical basis of the fuzzy snake as a new form of active contour, broadening the areas of application of this well known contour identification approach. It has been shown that the principal advantage of this new method, in parti-

cular, is the ability to exploit uncertain a *priori* knowledge, such as a verbal description of contour properties. The proposed new fuzzy energy functions present a level of abstraction which is higher and therefore closer to the human expert than that of algebraic energy functions. Parameters and weights with sometimes obscure meanings are replaced by the more intuitive linguistic interface provided by fuzzy logic expressions.

By allowing a more detailed object description it has been proved that the fuzzy snake approach can improve boundary detection in images of poor quality and also reduce computational complexity. Examples have been presented to verify the performance in being able to identify complex structures, most notably in the processing of MRI and X-ray based imaging sequences.

8 Acknowledgments

9 References

1 KASS, M., WITKIN, A., and TERZOPOULOS, D.: 'Snakes: Active contour models'. Proceedings of International conference on *Computer vision*, 1987, London, pp. 259–268
2 AMINI, A.A., TEHRANI, S., and WEYMOUTH, T.E.: 'Using dynamic programming for minimizing the energy of active contours in the presence of hard constraints'. Proceedings of 2nd international conference on *Computer vision*, 1988, pp. 95–99
3 OLSTAD, B.: 'Automatic wall motion detection in the left ventricle using ultrasonic images'. Proceedings of SPIE/SPSE conference on *Electronic imaging, Science & Technology*, San Jose, 1991
4 WILLIAMS, D., and SHAH, M.: 'A fast algorithm for active contours and curvature estimation', *Comput. Vis. Graph. Image Process.*, 1992, **55**, (1), pp. 14–26
5 McINERNEY, T., and TERZOPOULOS, D.: 'Deformable models in medical image analysis: a survey', *Med. Image Anal.*, 1996, **1**, (2), pp. 91–108
6 GRENANDER, U., and MILLER, M.: 'Representation of knowledge in complex systems', *J. R. Stat. Soc.*, 1993, **56**, pp. 249–503
7 DRYDEN, I., and MARDIA, K.V.: 'The statistical analysis of shape' (Wiley, London, 1998)
8 COOTES, T., TAYLOR, C., COOPER, D., and GRAHAM, J.: 'Active shape models – their training and application', *Comput. Vis. Image Underst.*, 1995, **61**, (1), pp. 38–59
9 OLSTAD, B.: 'Active contours with grammatical descriptions'. Proceedings of 6th international conference on *Image analysis and processing*, Como, Italy, September 1991
10 OLSTAD, B., and TORP, A.H.: 'Encoding of a priori information in active contour models', *Pattern Anal. Mach. Intell.*, 1996, **18**, (9), pp. 863–872
11 HÖWING, F., WERMSER, D., and DOOLEY, L.S.: 'Fuzzy snakes'. Proceedings of International conference on *Image processing and its applications*, Dublin, Ireland, July 1997, pp. 627–630
12 HÖWING, F., DOOLEY, L.S., and WERMSER, D.: 'Linguistic contour modelling through fuzzy snakes', *in* MOHAMMADIAN, M. (Ed.): 'Computational intelligence for modelling control and automation' (IOS Press, Amsterdam, 1999), pp.384–389
13 ZIMMERMANN, H.J.: 'Fuzzy set theory and its applications' (Kluwer, 1991)
14 BANDEMER, H.: 'Einführung in Fuzzy-Methoden' (Akademie-Verlag, Berlin, 1993)
15 KOSKO, B.: 'Neural networks and fuzzy systems' (Prentice-Hall, Englewood Cliffs, NJ, 1992)
16 TILLI, T.: 'Mustererkennung mit Fuzzy-Logik' (Franzis-Verlag, München, 1993)
17 BEZDEK, J.C. and PAL, S.K., (eds.): 'Fuzzy models for pattern recognition' (IEEE Press, New York, 1992)
18 ZADEH, L.A.: 'Fuzzy Sets', *Inf. Control.*, 1965, **8**, pp. 338–353
19 HÖWING, F., BÜLOW, H., WERMSER, D., DOOLEY, L.S., and THOMA, W.: 'Automatic motion analysis of bones from MR sequences'. Proceedings of the International conference on *Image processing and its applications*, Manchester, UK, July 1999, pp. 397–401
20 HÖWING, F., WERMSER, D., and DOOLEY, L.S.: 'Recognition and tracking of articulatory organs in X-ray image sequences', *Electron. Lett.*, 1996, **32**, pp. 444–445
21 HÖWING, F., DOOLEY, L.S., and WERMSER, D.: 'Tracking of non-rigid articulatory organs in X-ray image sequences', *Comput. Med. Imaging Graph.*, 1999, **23**, (2), pp. 19–27
22 MEYER, T.: 'Erarbeitung und Untersuchung von kantenorientierten Verfahren zur Vermessung der Elemente des Vokaltrakts in Röntgenbildern'. Master's thesis, FH Braunschweig/Wolfenbüttel, FB E, Germany, 1995
23 SCHALKOFF, R.J.: 'Digital image processing and computer vision' (Wiley, 1989)

10 Appendix: Algebraic energy functions

The internal energy used for all experiments was [20]

$$E_{int} = f(|k - m|) = e^{|k-m|} \tag{3}$$

The following external and constraint energy functions were proposed in [21]. The external energy computes an evidence value e for every candidate pixel $c_{n,m}$, denoting the possibility of it being a contour pixel:

$$E_{image}(c_{n,m}) = w_{edge}\|e_{edge}(c_{n,m})\| + w_{region}\|e_{region}(c_{n,m})\|$$
$$+ w_{motion}\|e_{motion}(c_{n,m})\| \tag{4}$$

$\| \|$: normalises the respective evidences to the range $[0, 1]$. High evidences correspond to low energy values.

Edge-based evidence:

$$e_{edge}(c_{n,m}) = \nabla_\perp(C_n, c_{n,m})\delta \tag{5}$$

$\nabla_\perp$: gradient magnitude (cf. [23]) along the search line C_n which is perpendicular to the initial contour hypothesis P, $\delta = 1$: rising edge, $\delta = -1$: falling edge.

Region-based evidence:

$$e_{region}(c_{n,m}) = -|\nabla(\Theta, c_{n,m})|$$

$$\Theta(x_i, y_j) = \begin{cases} 1 & : \quad R(x_i, y_j) \le \theta \\ 0 & : \quad R(x_i, y_j) > \theta \end{cases} \tag{6}$$

θ: grey level threshold to weaken the edge of dark occluding objects; depends on *a priori* knowledge. $R(x_i, y_j)$: a pixel from the region of interest R.

Motion-based evidence:

$$e_{motion}(c_{n,m}) = \epsilon_{med}(c_{n,m}),$$

$$\epsilon_{med} = \sqrt{\nabla(|R_{t-d_1} - R_t|)}\sqrt{\nabla(|R_t - R_{t+d_2}|)} \tag{7}$$

A moving edge detector (*med* [23]). R_i: region of interest extracted from images at instances $i = t$, $i = t - d_1$ and $i = t + d_2$, where d_1 and d_2 are constants.

For the energy component E_{con} several constraints can be applied. In this paper, a general form is introduced, where κ represents a measure of the constraining feature. This feature value is weighted by an assessment function v which is high when the actual feature value for a candidate vertex is similar to a prescribed value:

$$E_{con}(c_{n,m}) = \|v(\kappa(c_{n,m}))\| \tag{8}$$

$$\kappa(c_{n,m}) = \angle(c_{n-2,T(n-1,k)}, c_{n-1,k}, c_{n,m}) \tag{9}$$

$c_{n,m}$: the current vertex in the dynamic programming optimisation; $c_{n-1,k}$: its possible predecessors; $c_{n-2,T(n-1,k)}$: their optimal predecessors;

$$v(\gamma) = \begin{cases} 0 & : \quad \gamma^{min} > \gamma > \gamma^{max} \\ \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{\gamma-\bar{\gamma}}{\sigma})^2} & : \quad \gamma^{min} \le \gamma \le \gamma^{max} \end{cases} \tag{10}$$

$v(\gamma)$ is a Gaussian function, the parameters of which define either a convex or concave contour. $\bar{\gamma}$: desired local angle within an allowable range $[\gamma^{min}, \gamma^{max}]$. σ^2: certainty with which $\bar{\gamma}$ is known *a priori*.

Tracking of non-rigid articulatory organs in X-ray image sequences

Frank Höwing[a,*], Laurence S. Dooley[b], Diederich Wermser[c]

[a]*School of Electronics, University of Glamorgan, also with (current postal address): Fachbereich Elektrotechnik,
Fachhochschule Braunschweig/Wolfenbüttel, Salzdahlumer Str. 46/48, 38302 Wolfenbüttel, Germany*
[b]*School of Electronics, University of Glamorgan, Pontypridd, Mid Glamorgan CF37 1DL, UK*
[c]*Fachbereich Elektrotechnik, Fachhochschule Braunschweig/Wolfenbüttel, Salzdahlumer Str. 46/48, 38302 Wolfenbüttel, Germany*

Abstract

This article presents a system for the automated tracking of non-rigid anatomic structures in two-dimensional image sequences, which was primarily applied to X-ray image sequences of the vocal tract. In this particular application articulatory organs have to be measured to investigate the complex dynamic characteristics of human speech production. Of particular interest is a robust boundary detection of non-rigid organs such as lips and tongue. To solve this ill-posed detection problem under the presence of transparently superimposing structures, varying textural appearances of organs and noise, a two-level system is proposed. At the lower level, several edge-, region-, and motion-based image operators are combined to exploit their respective benefits and concomitantly compensate for their deficiencies. For the sake of precision, the result of these operators are not represented as larger tokens, such as line segments, but remain pixel-related cues or image evidences. At the higher level, an active contour-based component allows for the introduction of a priori knowledge about the object to be detected. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: X-ray image sequences; Vocal tract; Boundary detection; Non-rigid objects; Active contours

1. Introduction

The detection of the boundary of an object in an image can be an intractable problem owing to such extraneous effects as noise and the image projection process, so the problem is underconstrained and does not possess a unique solution. In the attempt to solve this problem, computer vision research has propounded many different approaches. Methods that are intrinsically able to handle variations in the boundary to be detected are of particular interest. They allow for the tracking of deformable objects in image sequences [1], as well as for identification of several similar specimens of rigid objects [2].

As a top-down method for detecting non-rigid objects, active contours [3] or snakes are a well-recognised approach. In applications where image scenes are complex however, that is ones which comprise many occluding or even transparently superimposing structures, the general constraints for smoothness may not lead to an exact result.

If the segmentation result is subsequently used to provide features for the classification of an object, or to qualitatively describe a scene, then an imprecisely extracted contour may still be adequate. This article conversely assumes that the class of an object is already known, but that the segmentation has to be precise, despite poor image quality. A good example of an application where these conditions occur is the tracking and measurement of anatomical structures in medical X-ray image sequences.

2. Application: analysis of the vocal tract

The following application helps to substantiate the characteristics of medical X-ray image sequences. It is a good example of where precise segmentation of object boundaries must be obtained from images which have low localised contrast and generally poor quality.

Articulatory phonetics is a branch of linguistics that is concerned with the very complex dynamic characteristics of the articulatory organs of the human vocal tract. Understanding their motion and interrelation is an important basis for understanding human speech production. Apart from being a contribution to basic research, this knowledge is also valuable in speech therapy [4–6].

* Corresponding author. Tel.: +49/5331/939-381; fax: +49/5331/939-118.

E-mail addresses: f.hoewing@fh-wolfenbuettel.de (F. Höwing), lsdooley@glamorgan.ac.uk (L.S. Dooley), wermser@fh-wolfenbuettel.de (D. Wermser)
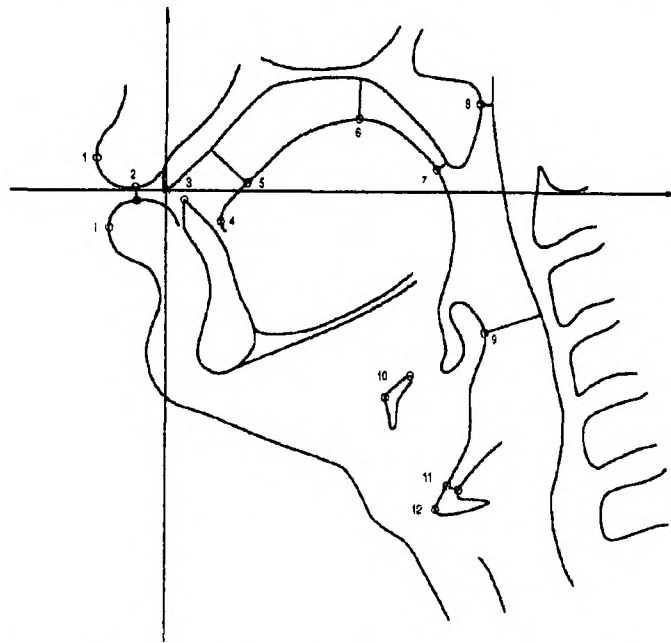
Fig. 1. Characteristic parameters of the articulatory organs. (1). Lip position, (2). Lip opening, (3). Front teeth, (4). Tip of tongue, (5). Front tongue, (6). Middle tongue, (7). Back tongue, (8). Velum, (9). Epiglottis, (10). Hyoid bone, (11). Glottal narrowing, (12). Glottis.

To analyse the articulatory organs it is necessary to determine the position of 12 characteristic parameters of the vocal tract defined on the midsagittal plane (cf. Fig. 1; for an anatomical background see Refs. [7–9]). For this purpose X-ray image sequences were taken, at a rate of 25 frames per second, from different speakers while uttering certain syllables. Fig. 2 shows a single frame from one such sequence.

For a statistically reliable analysis a very large number of images have to be processed, so manual measurement would be prohibitively time-consuming. However,

automated processing would importantly eliminate subjective influences from the measuring process.

2.1. Medical X-ray image sequences

There are two basic characteristics common to all X-ray images of human organs, which pose particular problems to their computer-based analysis:

- the images are blurry and have a low local contrast,
- despite the use of a contrast agent, relevant soft parts are occluded by more visible objects like bones, cartilage and teeth.

For these reasons, low-level image processing operations, like edge filters and thresholding, that are often successfully applied to industrial scenes are ineffectual.

The characteristics of the medical image sequences considered in this article may be summarised as follows (for X-ray image formation see Refs. [10–13]):

1. It is assumed that there is no explicit information on the specific X-ray imaging system. Information on the image formation process, other than general X-ray imaging characteristics (see below), is therefore not exploited. The lack of this knowledge is not so much a drawback for the development of a computer vision system, as it is an advantage for the generality of the solution and the range of other possible applications.

2. As X-rays penetrate the objects rather than being reflected from them, the resulting image is determined by the absorption coefficients of the objects.

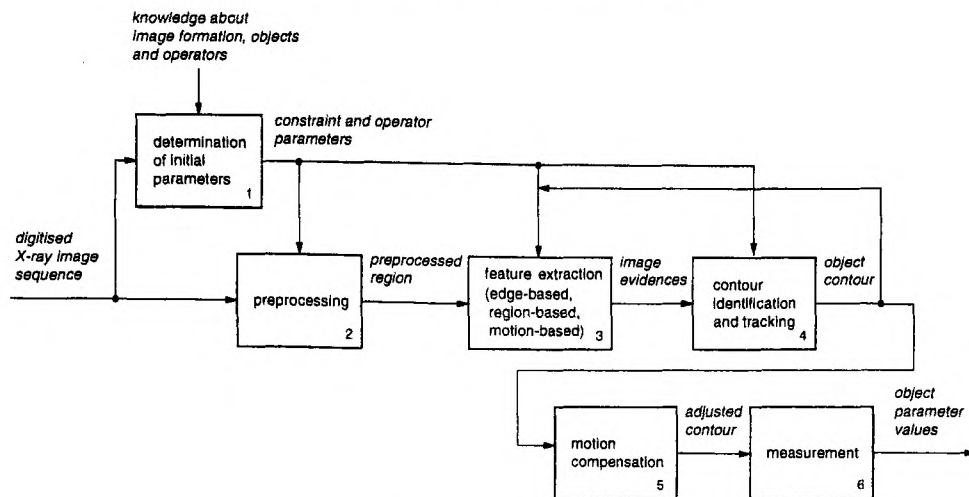Fig. 2. X-ray image of the human vocal tract.

Fig. 3. System structure.

3. An X-ray image is a two-dimensional, transparent projection of three-dimensional objects, so therefore there are superimposing structures in the image. Objects of a high density may even occlude other objects.

4. To make soft parts visible, a contrast agent may be applied. This substance changes the tissues' absorption coefficient and therefore their textural appearance in the image. The concentration of the contrast agent may vary (e.g., because of salivation in case of the earlier mentioned vocal tract images), so an object's appearance will in general be time varying.

5. It is assumed that no subtraction radiography is available, that is no images were taken without a probationer or, where applied, a contrast agent. This again widens the generality of the proposed system.

6. The scenes are complex. There are many structures of which only few may be of interest, that is there are many elements that may impair image processing.

7. Object boundaries are not sharply defined but rather blurry and have a low local contrast.

8. Local as well as global mean brightness may not be constant over time.

9. The image formation process may have introduced a considerable amount of noise.

3. Contour identification and measurement system

Fig. 3 shows the overall structure of the proposed modularised contour identification and tracking system, whose characteristics are:

- A sequence is processed frame by frame. Currently all objects are processed independently.
- Except for the first image of the sequence, where the relevant contours have to be marked by a human expert, the system is fully automatic.
- Multiple image evidences from different image operators are integrated.
- The detected contour is fed back into the system to improve detection in the next frame.
- As the probationer (or patient) may move relative to the camera, this motion is compensated for to obtain anatomically relevant measurement values.
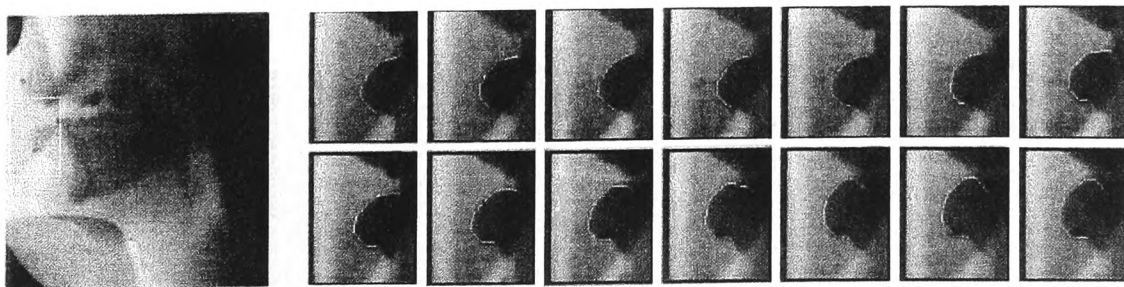

Fig. 4. Results of a dynamic-programming based contour tracing system using a minimum of a priori knowledge about the object. The images show the region of the lower lip in successive frames (from left to right and top to bottom), with the identified contour superimposed.
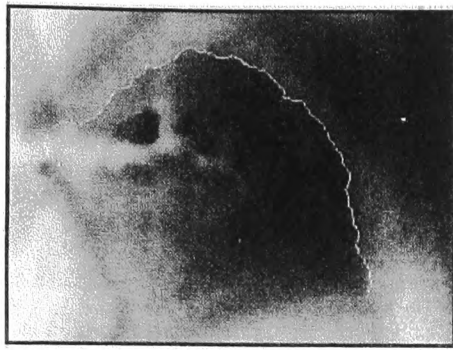
Fig. 5. Insufficient boundary detection of the tongue.

3.1. Integration of a priori knowledge

Explicit knowledge about image formation, object characteristics and image operator function is initially converted into system parameters by a human expert. For each image sequence, an initial contour hypothesis has to be drawn for each object. Based on this, the system is subsequently able to automatically process large image sequences without further interaction.

A simple contour tracing algorithm demonstrates the benefits resulting from the introduction of even the minimum of geometric a priori knowledge concerning the object to be detected. A dynamic programming algorithm described in [14,15] was applied to trace the contour of an object along a region of interest (ROI). The region is initially determined from a manually drawn contour hypothesis, introducing knowledge about the shape of the object to the system. The maximum distance that an object might travel between two frames determines the width of the ROI. Once a contour was determined it is used as a contour hypothesis for the next frame.

Unlike mere low-level based image processing techniques the algorithm is able to reliably identify the lips in a large number of images (see Fig. 4), so guaranteeing a closed and unique contour. The experiments showed, however, that such simple constraints are not always able to prevent the contour from being distorted by occluding objects. In situations where the object of interest is touching another object, like the other lip or the teeth, the contour traces an erroneous track.

Moreover the algorithm is not capable of identifying the boundary of the tongue. Fig. 5 shows that the contour tracing merely follows a maximum gradient path. Owing to the challenging image characteristics this path does not always correspond to the required object boundary.

These results emphasised the need for a better model, which also had to contain more specific geometrical constraints. With the proposed system these constraints were incorporated into an active contour based module, described in Section 3.4.

3.2. Preprocessing

The preprocessing module enhances the often considerably noisy and low-contrasted images. It also reduces the overall computational complexity by extracting a user-defined region of interest (ROI) for a particular object. (see Fig. 6).

An automatic image contrast enhancement is performed by computing an equalization interval based on the input image intensity frequency distribution. The interval is determined by searching the image histogram from both extrema until two values are found for which a specified percentage of the total image area falls between each value and the corresponding end of the histogram. This equalization is primarily applied to provide subsequent processing stages with a normalised intensity range, which facilitates parameter selection.

Noise reduction is obtained through standard median filtering of the ROI, exploiting the edge preserving properties of this operator. In our application a relatively large kernel size of 11×11 was used to reduce the large amount of noise introduced by the digitisation process.

As outlined in Section 2.1, details about the image formation process are not available. The operations used here are therefore general enough to make the preprocessing stage independent of a particular medical imaging device.

3.3. Feature extraction

The task of this module is to extract image features e_{image} that form the basis for the subsequent contour identification module. Unlike other approaches the proposed module does not segment tokens [16], such as geometric objects like lines or polygons, and either group or delete them afterwards
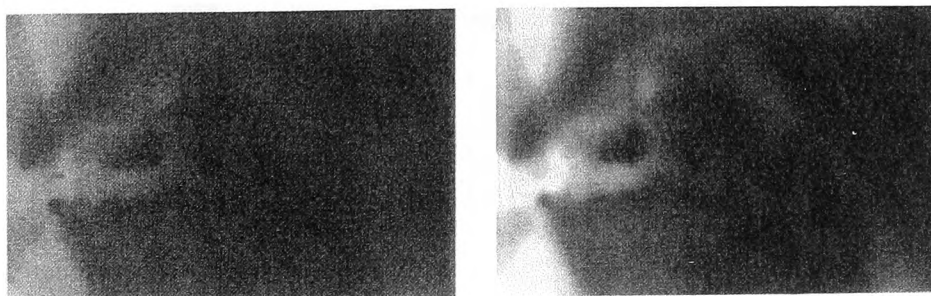

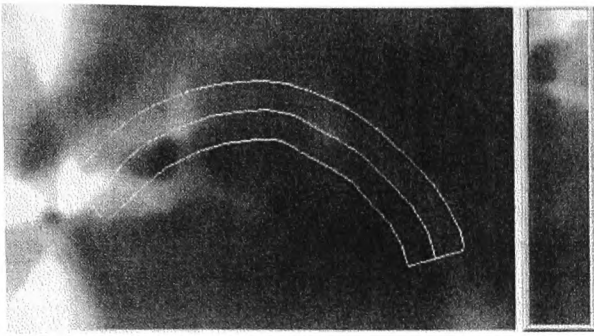Fig. 6. Region of interest for the tongue before and after preprocessing.

Fig. 7. The search space R is derived from an initial contour hypothesis (centre line).

based on their individual features. Instead, the features remain pixel-related in order to be able to identify a precise contour.

To exploit the benefits of the many existing image processing operators and at the same time compensate for their individual deficiencies, a structure is introduced that combines several edge-, region-, and motion-based low-level operators in a computationally efficient manner by a normalised weighted sum Eq. (1).

$$e_{image}(c) = w_{edge}\|e_{edge}(c)\| + w_{region}\|e_{region}(c)\|$$

$$+ w_{motion}\|e_{motion}(c)\| \quad (1)$$

The result of each operator, normalised to the range [0,1] through the norm $\|\cdot\|$, is regarded as an indication of to what degree a certain pixel c may belong to the object's contour. This local feature value is therefore referred to as image evidence e. It is also possible to incorporate operators that are able to deliberately detect structures which do not belong to the desired contour.

Experiments show that each operator may wrongly lead to low evidence values for contour pixels and/or high values for non-contour pixels. The weighted combination of all

evidences for a certain pixel compensates for the effect and leads to an attenuation of evidences of most contour pixels.

The individual evidence functions are very-much application dependent. The following definitions in Eqs. (2) to (4) have successfully been applied to the application presented in this article. Figs. 7 and 8 demonstrate the individual sub-components of e_{image} which are now discussed.

The edge-based evidence is given by

$$e_{edge}(c) = \nabla_{\perp}(c)\delta \quad (2)$$

where $\nabla_{\perp}$ is the gradient normal to the initial contour hypothesis. The factor δ specifies either a rising ($\delta = 1$) or falling edge ($\delta = -1$) respectively.

The region-based evidence given in Eq. (3) weakens the influence of occluding objects with grey levels below a prescribed threshold θ. Pixels within these objects are not influenced.

$$e_{region}(c) = -|\nabla(\Theta, c)|$$

$$\Theta(x_i, y_j) = 1 \;:\; R(x_i, y_j) \leq \theta \quad (3)$$

$$\Theta(x_i, y_j) = 0 \;:\; R(x_i, y_j) > \theta$$

Here $R(x_i, y_j)$ denotes a pixel from the region of interest $\mathbf{R}$ around the contour hypothesis (cf. Fig. 7). θ depends on a priori knowledge.

To exploit information from motion in the image sequence the third evidence value in Eq. (4) is applied.

$$e_{motion}(c) = \varepsilon_{med}(c),$$

$$\varepsilon_{med} = \sqrt{\nabla(|R_{t-d_1} - R_t|)}\sqrt{\nabla(|R_t - R_{t+d_2}|)} \quad (4)$$

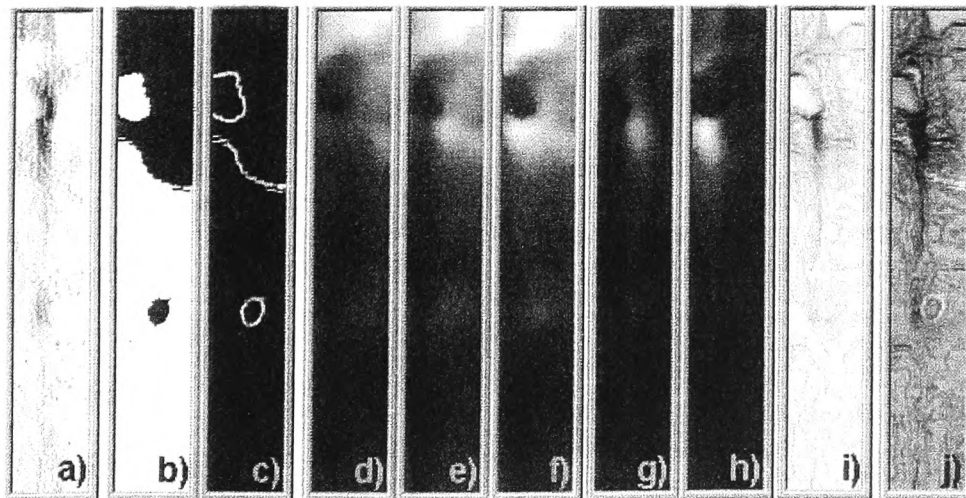This moving edge detector which is indicated by the subscript med [7], gives high evidence values for moving

Fig. 8. Sub-components of e_{image}, (a) e_{edge}, (b)$\Theta(\theta = 80)$, (c) e_{region}, (d) R_{t-d_1}, (e) R_t, (f) R_{t+d_2}, (g)$|R_{t-d_1} - R_t|$, (h) $|R_t - R_{t+d_2}|$, (i) e_{motion}, (j) e_{image} ($w_{edge} = 0.1$, $w_{region} = 0.1$, $w_{motion} = 0.8$).
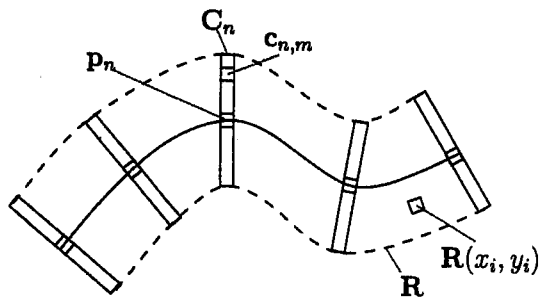
Fig. 9. Polygonal contour representation.

edges by multiplying the gradient of two difference images. The operation is performed on the same region of interest R_i extracted from images at instances $i = t$, $i = t - d_1$, and $i = t + d_2$, where d_1 and d_2 are constants.

3.4. Contour identification and tracking

The module described in this section represents the core of the proposed system. Its task is to identify the contour pixels of an object based on the evidences provided by the preceding modules described earlier.

A contour identification approach which not only uses an intermediate symbolic representation of an object, but also provides a method that links this representation with image features and a priori knowledge, is that of active contours or snakes. Active contours also allow for interaction with higher-level processes and may therefore represent a basis for further extensions of the proposed system.

Active contours were first introduced by Kass et al. [3]. They can be described as energy minimising splines or polygons. For each image, the algorithm needs an initial polygon $P = (p_0, p_1, ..., p_{N-1})$ consisting of N vertices $p_i = (x_i, y_i)$, where x_i and y_i are the spatial co-ordinates of p_i. The detected boundary is represented by the polygon $Q = (q_0, q_1, ..., q_{N-1})$ with $q_i = (x_i, y_i)$. Each q_i is selected from a set of candidates $C_i = (c_{i,0}, c_{i,1}, ..., c_{i,M-1})$ where in our application, the candidates $c_{i,j} = (x_j, y_j)$ are uniformly sampled along a search line normal to the initial polygon and intersecting p_i (cf. Fig. 9). Tracking of the contour is realised by processing a sequence, frame by frame and taking the resulting $Q(t)$ as the initial, estimated contour $P(t + 1)$ of the next frame.

With the polygonal representation it is possible to formulate an appropriate energy function E_{snake} for the object which is then minimised to obtain the desired contour Q. Moreover, the representation incorporates the simple connectivity of the objects and drastically reduces search space.

Kass et al. [3] proposed that an energy function be composed of the following three components:

- E_{int} represents the internal energy of the active contour, forcing it to act like a membrane or thin plate, producing a smooth shape.

- E_{image} represents an external force that guides the contour towards features in the image.
- E_{con} represents external constraint forces.

To minimise the energy function, the original algorithm of Kass et al. [3] involves four steps:

1. setting up a variational integral on the continuous plane
2. deriving a pair of Euler equations
3. discretisation
4. solving the discrete equations iteratively until convergence.

There are a number of problems resulting from this algorithm that were recognised, such as that the energy function must be a differentiable function which therefore constrains the range of possible models [17]. Also the vertices may move along the contour and cluster, because it is not possible to incorporate "hard constraints" such as a minimum inter-vertex distance [18]. As with most iterative approaches the problem of convergence is one of the main subjects in the literature (cf. Ref. [19] for an overview).

To overcome these problems, Amini [18] establishes the problem of energy minimisation as a discrete multi-stage decision process, which enables him to use a discrete dynamic programming algorithm to find an optimal solution. The results obtained by this method, however, still did not yield the precision needed in the processing of medical X-ray image sequences. Even after 30 iterations (convergence) the contour was not able to precisely identify the entire boundary of a non-occluded object [18].

This result is neither a fundamental drawback of the active contour approach nor of the dynamic programming optimisation – similar algorithms were applied successfully to medical images by Olstad (ultrasonic image sequence of the left ventricle [20]) and Geiger (angiograms and nuclear magnetic resonance (NMR) images of the left ventricle [21]). The reason is rather that it is necessary to fully exploit the potential of the approach by formulating the energy function according to the characteristics of image and object. In the proposed system, this was achieved by:

- introducing multiple image evidences from several sources to form E_{image},
- introducing an object specific constraint (E_{con}), that restricts the possible shapes of the active contour beyond a mere, general demand for smoothness.

The energy function of the active contour is minimised by the dynamic programming algorithm where S (n, m) represents the minimal energy level that is possible for the vertices $0, ..., n$ if the nth vertex is the candidate $c_{n,m}$. T (n, m) holds the index k ($k = 0, ..., M - 1$) that minimises the expression in line 8 and thus points to the optimal predecessor of the candidate $c_{n,m}$. After all vertices were processed, the new boundary is obtained by tracing back the pointers, beginning with the candidate that has a minimal S (N − 1, m) value.

Dynamic programming algorithm

1. **for all** m
2. $\quad S(0, m) = w_{image}E_{image}(\mathbf{c}_{0,m})$
3. **for all** m
4. $\quad S(1, m) = \min_k [w_{image}E_{image}(c_{1,m}) + w_{int}E_{int} + S(0, k)]$
5. $\quad T(1, m) = k^{min}$
6. **for** $n = 2, ..., N - 1$
7. $\quad$ **for** $m = 0, ..., M - 1$
8. $\quad\quad S(n, m) = \min_k [w_{image}E_{image}(c_{n,m}) + w_{int}E_{int} + w_{con}E_{con} + S(n - 1, k)]$
9. $\quad\quad T(\text{n}, m) = \text{k}^{min}$

The weighting factors w_{image}, w_{int} and w_{con} control the relative influences of the energy function's components.

In our application the image energy is derived from the image evidences described in Section 3.3: (see Eq. (5)

$$E_{image} = 1 - \mathbf{e}_{image} \tag{5}$$

The internal energy of the active contour is denoted [20] as

$$E_{int} = \|e^{|k-m|}\|, \tag{6}$$

favouring candidates with a curvature similar to the initial hypothesis. Therefore an inter-vertex force counteracting the external force is applied, smoothing the contour. The function $\|\cdot\|$ normalises E_{int} to the range [0,1], where 0 indicates a strong similarity, to facilitate parameter settings (see Eq. (6)).

A geometrical constraint presented in [22] is applied, which explicitly introduces knowledge on the expected shape of the contour and which can be easily incorporated into the dynamic programming algorithm. Eq. (7) favours convex or concave contours by weighting the angle γ which is measured in the open polygon consisting of the current candidate vertex $\mathbf{c}_{n,m}$, its possible predecessor $\mathbf{c}_{n-1,k}$, and the optimal predecessor of the latter, $\mathbf{c}_{n-2,T(n-1,k)}$. By always measuring γ in the same direction, setting the possible range $[\gamma^{min}, \gamma^{max}]$, the expected angle $\bar{\gamma}$ and the standard deviation σ, the resulting contour can be forced to bend in a desired direction.

$$E_{con} = \|\nu(\gamma)\|,$$

$$\nu(\gamma) = 0 \ : \ \gamma^{min} > \gamma > \gamma^{max}$$

$$\nu(\gamma) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{\gamma-\bar{\gamma}}{\sigma}\right)^2} \ : \ \gamma^{min} \le \gamma \le \gamma^{max} \tag{7}$$

The use of a Gaussian weighting function (again normalised to the range [0,1] through a normalisation function $\|\cdot\|$) allows exploitation of uncertain knowledge concerning the actual shape of the object. For instance if σ is small then the tolerance range around $\bar{\gamma}$ where E_{con} gives a good assessment is narrower than for higher values of σ. In theory, the weighting function applied in Eq. (7) could be replaced by any mathematical distribution which exhibited the appropriate properties.

4. Results

Figs. 10 and 11 demonstrate the performance of the proposed system, showing some results of identifying the tongue in a number of frames of two X-ray sequences. Both
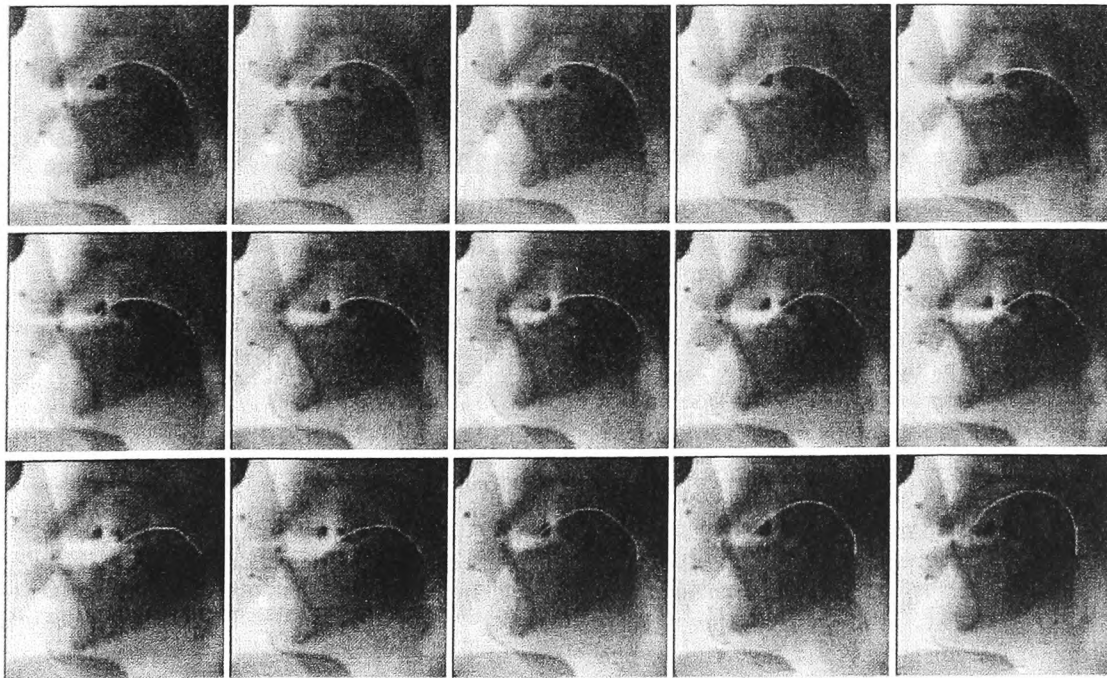

Fig. 10. Results of the proposed system identifying the tongue. Every second frame of a sequence is shown.

Fig. 11. Results of the tongue tracking system on a different probationer. Every third frame of a sequence is shown.

sequences of different probationers were successfully processed using the same parameter settings. The weighting factors of the normalised energy components were $w_{image} = 0.6$, $w_{int} = 0.15$ and $w_{con} = 0.25$. The constraint parameters were set to $\bar{\gamma} = \pi - 0.2$, $\sigma = 0.2$, $\gamma^{min} = \pi - 0.4$ and $\gamma^{max} = \pi$. These settings were easily found as small changes did not affect the results. The parameters of the geometrical constraint proved to be particularly robust.

The tongue is partially occluded by teeth, cheek bones and fillings. As a result of salivation the distribution of the contrast agent which was applied orally, is not constant. Although these influences lead to varying features of the boundary, the tongue is detected correctly. Even when the back tongue touches the velum so there is little distinction to be made between the two articulators, the active contour correctly aligns with the boundary of the tongue. Fig. 11 demonstrates the same robustness when the middle tongue touches the soft palate. This result is of particular importance to the analysis of speech production, as the location and diameter of constrictions constitute important parameters to characterise a particular sound.

The final image in Fig. 10 shows a concave section in the front tongue, with the boundary correctly aligned with salient image features. This demonstrates the ability of the system to locally violate the geometrical constraint under the presence of strong image evidences.

5. Conclusions

The accurate detection and tracking of boundaries of non-rigid objects still pose many problems to computer vision systems. Image material such as medical X-ray image sequences involving transparently superimposing structures and varying textures are particularly challenging. To solve the underconstrained detection problem, a system was presented which combines the results of several low-level image operators to form pixel related image evidences. This method compensates for the individual deficiencies of the low-level operators, so the segmentation process does not rely on precise and unique low-level processing. Also unlike traditional techniques, neither grouping nor deletion of low-level tokens takes place, thus eliminating the propagation of potentially erroneous segmentations to a higher processing level.

As an example application the detection and tracking of

the lips and tongue in sagittal X-ray image sequences of the human vocal tract was presented. The successful boundary detection, which requires only an initial user interaction, forms the basis for a subsequent measuring of characteristic parameters of the articulatory organs. This automated process replaces the time consuming and inaccurate manual measurement of a very large number of X-ray images, to facilitate a statistical analysis.

Within the area of application considered in this article, a boundary detection cannot be successfully achieved without a model of the object to be detected. The proposed system integrates such a priori knowledge into an active contour based approach. Although the object model itself remains comparatively simple, the complete system is able to successfully identify the contours of simply connected deformable objects. Extensive experiments involving vocal tract image sequences have shown, that in most cases the best segmentation result is achieved in a single iteration, so problems related to the convergence of the algorithm and to possible clustering of contour nodes were overcome.

The system has also been proven to be robust against small changes in its weighting and constraint parameters and is therefore easily adaptable to similar tasks.

References

[1] Wermser D, Höwing F. Automatische Kennzeichnung und Vermessung in Röntgenbildsequenzen. Biomedical Journal, 1996:17–19.

[2] Lipson P, Yuille A, O'Keeffe D, Cavanaugh J, Taaffe J, Rosenthal D. Deformable templates for feature extraction from medical images. In: Faugeras O, editor. Computer vision-ECCV 90: First European Conference on computer Vision, Springer, 1990. pp. 413.

[3] Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models. In Proceedings of International Conference on Computer Vision, (London), 1987, pp. 259–268, June.

[4] Metoui M. Software-Einsatz für Forschung und Lehre in der artikulatorischen Phonetik. In: Dette K, Haupt D, editors. Multimedia und Computeranwendungen in der Lehre, Springer, 1992.

[5] Metoui M. Coarticulation: Statistical feature analysis with computer-aided processing of X-ray tracings. In 6. Expertentreffen des Arbeitskreises Pädagogische Software mit digitaler Sprachverarbeitung, Technische Universität Braunschweig, 1993.

[6] Metoui M. Phonetische Datenverarbeitung mit S.P.A.T. Möglichkeiten, Folgen, Herausforderungen, Forschungsmagazin der Universität Mainz, Sonderheft, 1994, pp. 18–25.

[7] Lippert H. Anatomie. 5. Text und Atlas. München, Wien, Baltimore: Urban and Schwarzenberg, 1989.

[8] Makroskopische und mikroskopische Anatomie des Menschen. In:

Benninghoff A, Straubesand J, editors. 14. Cytologie, Histologie, allgemeine Anatomie und Anatomie des Bewegungsapparates, 1. München; Wien; Baltimore: Urban und Schwarzenberg, 1985.

[9] Wörner H. Röntgen beim Zahnarzt. Köln. Deutscher Ärzte-Verlag, 1990.

[10] Erfassung und maschinelle Verarbeitung von Bilddaten. In: Kazmierczak H, editor. Berlin: Springer, 1980.

[11] M. Fröder, Darstellung geringkontrastiger Objekte im menschlichen Schädel mit rechnergestützter Röntgenvideotechnik. PhD thesis, Universität Erlangen-Nürnberg, Technische Fakultät, 1986.

[12] Cho Z, Jones JP, Singh M. Foundations of medical imaging. New York: Wiley, 1993.

[13] Rupp S. Digitale Radiographie-Optimierung der Bildqualität durch Bildverarbeitung. Düsseldorf. VDI-Verlag, 1991.

[14] Gerbrands JJ. Segmentation of noisy images. PhD thesis, Technical University of Delft, 1988.

[15] Bässmann H, Besslich PW. Konturorientierte Verfahren in der digitalen Bildverarbeitung. Berlin: Springer, 1989.

[16] Pinz A. Bildverstehen. Berlin: Springer, 1994.

[17] Storvik G. A bayesian approach to dynamic contours through stochastic sampling and simulated annealing. IEEE Transactions on Pattern Analysis and Machine Intelligence 1994;16:976–986.

[18] Amini AA, Tehrani S, Weymouth TE. Using dynamic programming for minimizing the energy of active contours in the presence of hard constraints, in Proceedings of the 2nd International Conference on Computer Vision, 1988, pp. 95–99.

[19] Leymarie F, Levine MD. Tracking deformable objects in the plane using an active contour model. IEEE Transactions on Pattern Analysis and Machine Intelligence 1993;15:617–634.

[20] Olstad B. Automatic wall motion detection in the left vertricle using ultrasonic images, in Proceedings of SPIE/SPSE: Electronic imaging. Science and Technology, (San Jose), 1991.

[21] Geiger D, Gupta A, Costa LA, Vlontzos J. Dynamic programming for detecting, tracking and matching deformable contours. IEEE Transactions on Pattern Analysis and Machine Intelligence 1995;17:294–302.

[22] Höwing F, Wermser D, Dooley L. Recognition and tracking of articulatory organs in X-ray image sequences. IEE Electronics Letters 1996;32:444–445.

Frank Höwing received his Dipl.-Ing. Degree in Computer Science from the Fachhochschule Braunschweig/Wolfenbüttel in 1994. After working on an industrial project involving neural networks design for vehicle transmission control, he rejoined the Fachhochschule, where he is currently a scientific assistant in the Department of Electrical Engineering. Since 1995 he was a registered Ph.D. research student at the University of Glamorgan.
His research interests are in the fields of computer vision and fuzzy logic.

Laurence Dooley received his B.Sc., M.Sc., Ph.D. Degrees in Electrical Engineering from the University College of Wales, Swansea in 1981, 1983 and 1987 respectively. He worked as a consultant for industry in the area of digital sound synthesiser design for Marine Simulator Systems, before joining the Department of Electronics and Information Technology at the University of Glamorgan in 1986, where he is presently a Senior Lecturer and heads the Media Technology Group. He has lectured extensively around the world and in 1989 was a Visiting Scholar in the Department of Electrical Engineering at the University of Sydney, Australia.
His principal research interests are in the fields of digital signal and image processing, video technology, wavelets and medical imaging. He is a Senior Member of the IEEE, a Chartered Engineer (C. Eng), a corporate member of the British Computer Society and the Society of Computer Simulation.

Diederich Wermser received his Dipl.-Ing. and Dr.-Ing. Degrees in Electrical Engineering from the University of Hannover in 1980 and 1984 respectively. Between 1984 and 1988 he joined Scientific Control Systems, a Hamburg based research company, being responsible for several projects in the areas of knowledge based systems and distributed systems. Between 1989 and 1993 he was leading the telecommunications projects and strategic studies department of NTE NeuTech, a Munich based research subsidiary of the Axel Springer publishing company in Germany. Since 1994 he has been a Professor at the Fachhochschule Braunschweig/Wolfenbüttel and head of the telecommunications laboratory.
His research interests are in the field of digital image processing and telecommunication systems.

Recognition and tracking of articulatory organs in X-ray image sequences

F. Höwing, D. Wermser and L.S. Dooley

Deformable organs of the human vocal tract are detected. An active contour is optimised by a dynamic programming algorithm, for which a new constraint is presented that introduces a priori knowledge on the shape of the expected boundary. The algorithm is able to detect convex and concave objects even when the image quality is poor.

Introduction: The analysis of the articulatory organs of the human vocal tract is important in understanding and modelling speech. A comprehensive model of human speech production can be derived [1] from the precise measurement of the motion of characteristic points in the vocal tract, with X-ray image sequences being used to provide information upon all organs. As a very large number of images have to be measured, an accurate and robust automated system is required.

This Letter presents a central element of an image processing system, capable of recognising and tracking deformable organs in large X-ray image sequences. Most of the problems that have been encountered and subsequently overcome were due to features of the X-ray images, such as low contrast, varying textures and superimposing objects.

Algorithm: The detection is based on the concept of Active Contours, which can be described as energy minimising splines or polygons [2]. For each image, the algorithm needs an initial polygon $\mathbf{P} = (\mathbf{p}_0, \mathbf{p}_1, ..., \mathbf{p}_{N-1})$ consisting of N vertices $\mathbf{p}_i = (x_i, y_i)$, where x_i and y_i are the spatial co-ordinates of $\mathbf{p}_i$. The detected boundary is represented by the polygon $\mathbf{Q} = (\mathbf{q}_0, \mathbf{q}_1, ..., \mathbf{q}_{N-1})$ with $\mathbf{q}_i = (x_i, y_i)$. Each $\mathbf{q}_i$ is selected from a set of candidates $\mathbf{C}_i = (\mathbf{c}_{i,0}, \mathbf{c}_{i,1}, ..., \mathbf{c}_{i,M-1})$, where the candidates $\mathbf{c}_{i,j} = (x_j, y_j)$ are uniformly sampled along a search line normal to the initial polygon and intersecting $\mathbf{p}_i$. Tracking of the contour is realised by processing a sequence, frame by frame and taking the resulting $\mathbf{Q}(t)$ as the initial, estimated contour $\mathbf{P}(t + 1)$ of the next frame.

The energy function of the active contour is minimised by the dynamic programming algorithm in Table 1. $S(n,m)$ represents the minimal energy level that is possible for the vertices 0, ..., n if the nth vertex is the candidate $\mathbf{c}_{n,m}$. $T(n,m)$ holds the index k ($k = 0, ..., M-1$) that minimises the expression in line 8 and thus points to the optimal predecessor of the candidate $\mathbf{c}_{n,m}$. After all vertices have been processed, the new boundary is obtained by tracing back the pointers, beginning with the candidate that has a minimal $S(N-1, m)$ value.

U_1 represents an external force that guides the contour towards features in the image. In our particular application the gradient of

Table 1: Dynamic programming algorithm

1.	for all m
2.	$S(0,m) = \alpha_1 U_1(\mathbf{c}_{0,m})$
3.	for all m
4.	$S(1,m) = \min_k[\alpha_1 U_1(\mathbf{c}_{1,m}) + \alpha_2 e^{|k-m|} + S(0,k)]$
5.	$T(1,m) = k^{min}$
6.	for $n = 2 ... N-1$
7.	for $m = 0 ... M-1$
8.	$S(n,m) = \min_k[\alpha_1 U_1(\mathbf{c}_{n,m}) + \alpha_2 U_2 + S(n-1,k)]$
9.	$T(n,m) = k^{min}$

the image intensities along the search line is employed. The second component U_2 represents an internal force that restricts the possible shapes of the contour. The weighting factors α_1 and α_2 control the relative influence of the two components. In eqn. 1, a constraint is introduced that forces the active contour to preserve a curvature similar to that of the initial estimate,

$$U_2 = e^{|k-m|} \qquad (1)$$

With this constraint the active contour is already able to move to small areas of weak image features when necessary to produce an overall smooth contour. Where these areas are too large, however, or where strong features occur that do not belong to the object boundary, this constraint is insufficient.

A geometrical constraint is therefore proposed, which explicitly introduces a priori knowledge on the expected shape of the contour and which can be easily incorporated into the dynamic programming algorithm. Eqn. 2 favours convex or concave contours by weighting the angle γ which is measured in the open polygon consisting of the current candidate vertex $\mathbf{c}_{n,m}$, its possible predecessor $\mathbf{c}_{n-1,k}$, and the optimal predecessor of the latter, $\mathbf{c}_{n-2,T(n-1,k)}$. By always measuring γ in the same direction, setting the possible range $[\gamma^{min}, \gamma^{max}]$, the expected angle $\bar{\gamma}$ and the standard deviation σ, we can force the resulting contour to bend in a desired direction,

$$U_2 = \begin{cases} 0 & \gamma^{min} > \gamma > \gamma^{max} \\ -\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{\gamma-\bar{\gamma}}{\sigma}\right)^2} & \gamma^{min} \le \gamma \le \gamma^{max} \end{cases} \qquad (2)$$

Results: The algorithm has been applied succesfully to detect the boundaries of the lips and part of the tongue in large X-ray image sequences, an example of which is shown in Fig. 1a. It has proven to be especially robust against changing image properties.
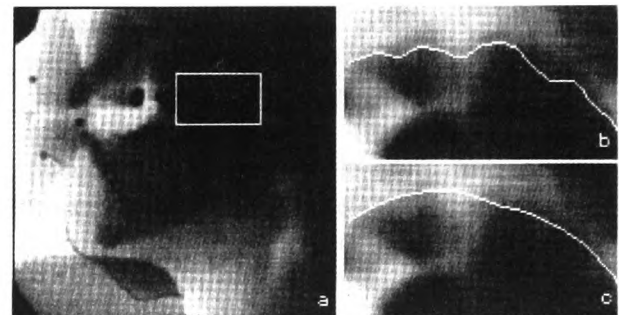
Fig. 1 X-ray image of the vocal tract

a Vocal tract
b Contour detection without constraint
c Contour detection with proposed shape constraint

Fig. 1b shows an image where a portion of the back tongue has been detected without any constraint. It is clear that the contour is neither very smooth nor does it follow the desired object boundary particularly well. The contour has just been dragged towards the strongest image features within the search area.

Application of the proposed constraint in eqn. 2 gives to the boundary in Fig. 1c. It can be seen that the active contour is now able to bridge areas of weak image feature very well, avoiding misleading strong features to produce a contour of the desired shape.

Conclusion: A new shape constraint for active contours based on a dynamic programming algorithm is proposed. It allows for a robust and precise detection of convex or concave object boundaries in X-ray image sequences with low contrast and varying textures. The new constraint is insensitive to small changes in its parameters and is therefore easily applicable.

F. Höwing, and L.S. Dooley (Department of Electronic and Information Technology, University of Glamorgan, Pontypridd, Mid Glamorgan, CF37 1DL, United Kingdom)

D. Wermser (Fachbereich Elektrotechnik, Fachhochschule Braunschweig/Wolfenbüttel, Salzdahlumer Str. 46/48, 38302 Wolfenbüttel, Germany)

F. Höwing is also with Fachbereich Elektrotechnik, Fachhochschule Braunschweig/Wolfenbüttel, Salzdahlumer Str. 46/48, 38302 Wolfenbüttel, Germany

References

1 METOUI, M.: 'Coarticulation: Statistical feature analysis with computer-aided processing of X-ray tracings'; 6. Expertentreffen des Arbeitskreises Pädagogische Software mit digitaler Sprachverarbeitung. Technische Universität Braunschweig, Germany, 1993

2 OLSTAD, B.: 'Automatic wall motion detection in the left ventricle using ultrasonic images'. SPIE/SPSE symposium on electronic imaging. Science & Technology, San Jose, 1991

Automatic Motion Analysis of Bones from MR Sequences

F. Höwing[*,+], H. Bülow[*,+], D. Wermser[*], L. Dooley[+], W. Thoma[‡]

[*] Fachhochschule Braunschweig/Wolfenbüttel, FB E
Salzdahlumer Str. 46/48, 38302 Wolfenbüttel
Email: {f.hoewinglh.buelowld.wermser}@fh-wolfenbuettel.de

[+] University of Glamorgan, School of Electronics,
Pontypridd, Mid Glamorgan, CD37 1DL, UK
Email: lsdooley@glamorgan.ac.uk

[‡] Orthopädische Universitätsklinik Frankfurt/Main, Stiftung Friedrichsheim
Marienburgstr. 2, 60582 Frankfurt/Main

Abstract. In many cases articular damages cannot be diagnosed through an examination of a single image. A motion analysis of a joint's bones might be necessary to make a reliable diagnosis [1,2]. Examples are lesions of the ligaments and cartilage of the knee or in the cervical and lumbar regions of the vertebral. This paper presents a novel system to diagnose lesions of the ligaments of the wrist (carpal instabilities [3]). The method is particularly well-suited to aid in the diagnosis of the *scapho-lunate instability*. This damage is a common injury after accidents involving the wrist. The lesion occurs when the ligaments between the Scaphoid and the Lunate are torn [4]. Motion graphs (Fig. 4) show the rotation as well as the translation of the carpal bones. The measurement is performed relative to an anatomic co-ordinate system defined by the distal end of the Radius. Compared to other applications [5] a motion analysis of wrist bones is more difficult because there are many bones with a similar shape which complicates their identification. Furthermore some of the bones may tilt, that is they may rotate around axes not perpendicular to the view plane. This results in a varying appearance of the bones in the sliced magnetic resonance (MR) images.

1 Problem

Availability of Nuclear Magnetic Resonance Imaging allows scanning of entire sequences of images of bones and joints without harmful dosage of radiation. Analysis of such sequences allows a much more reliable diagnosis of lesions of the ligaments compared to methods in use today such as single x-ray images [1,2,3]. However, a necessary scanning procedure with a sufficient number of positions requires approximately 100 2-D images for every patient. The manual evaluation of such a number of images in the daily medical diagnostic is not feasible. The aim of the proposed approach here is the automatic processing of these images in order to obtain motion graphs which allow an easy medical diagnostic. For the recognition of lesions of the ligaments the representation of translation and rotation of the carpal bones with respect to a coordinate system defined by the radius proved to be most suitable. Using cadaveric specimen such investigations have been carried out by implantation of markers [4]. Because of the considerable exposure to radiation, in vivo analysis of such movements has been carried out only with very coarse resolution. The method proposed here allows for a much finer resolution of the bone movement determined (Fig. 4).

Compared to the motion analysis of the knee or the spine the measurement of the motion of carpal bones is much more complicated. This is because of the number of bones with similar appearance which complicates the identification of the bones. Furthermore a tilt of some of the carpal bones is responsible for a change in appearance within the NMR slices during the sequence.

2 Automatic measurement

The overall system comprises the following components:

- Image acquisition – Depending on the flexibility of each patient approximately 8 different positions of the wrist will be scanned. For each wrist position 12 layers of the hand are acquired.

- Layer selection – An approach based on the Fourier-Mellin transform [6] allows for the selection of an MR layer which is best suitable for the measurement by comparing the input layers with a reference image (Fig. 1).

- Segmentation – An adaptive threshold is applied to an automatically selected region of interest (ROI). To obtain a higher precision the algorithm is applied in two stages to the ROI of the wrist and to smaller ROIs of the individual bones.

- Identification of the bones – Constrained by their possible motion the relevant bones are identified through an analysis of the shape and position of a set of candidate bones (Fig. 2).

- Measurement of translation and rotation – For each bone its major axis and centroid is determined (Fig. 3).

- Motion graphs – The measurement results of usually about 8 different positions of the wrist are collected (Fig. 4). The coordinate system which is taken as reference will be determined by a concave curvature of the radius in order to compensate unavoidable movement of the patients arm.

The performance of the system is demonstrated by automatically measuring the motions of the bones of 158 wrist positions of 20 patients to date. A resulting number of 1106 bones were identified correctly. The segmentation was highly successful for the most relevant bones Scaphoid, Lunate and Radius (Tab. 1). A good segmentation was also obtained for other carpal bones, allowing the system to be applied to the diagnosis of other carpal instabilities as well.

Bone	Correct segmentation
os hamatum	77,8%
os capitatum	94,9%
os trapezoideum	90,5%
os triquetrum	89,2%
os lunatum	94,3%
os scaphoideum	96,8%
Radius	97,5%

Table 1: Success rate for automatic segmentation of the carpal bones in 158 wrist positions.

3 Clinical use

In order to ease the introduction of this system to daily use in medical diagnosis, an interactive graphical user interface is under development [7]. The intermediate results of the different processing steps are automatically tested for plausibility. If errors are detected, the user is automatically requested for a manual correction. A description of the appearance of specific lesions in motion graphs is presented in [8].
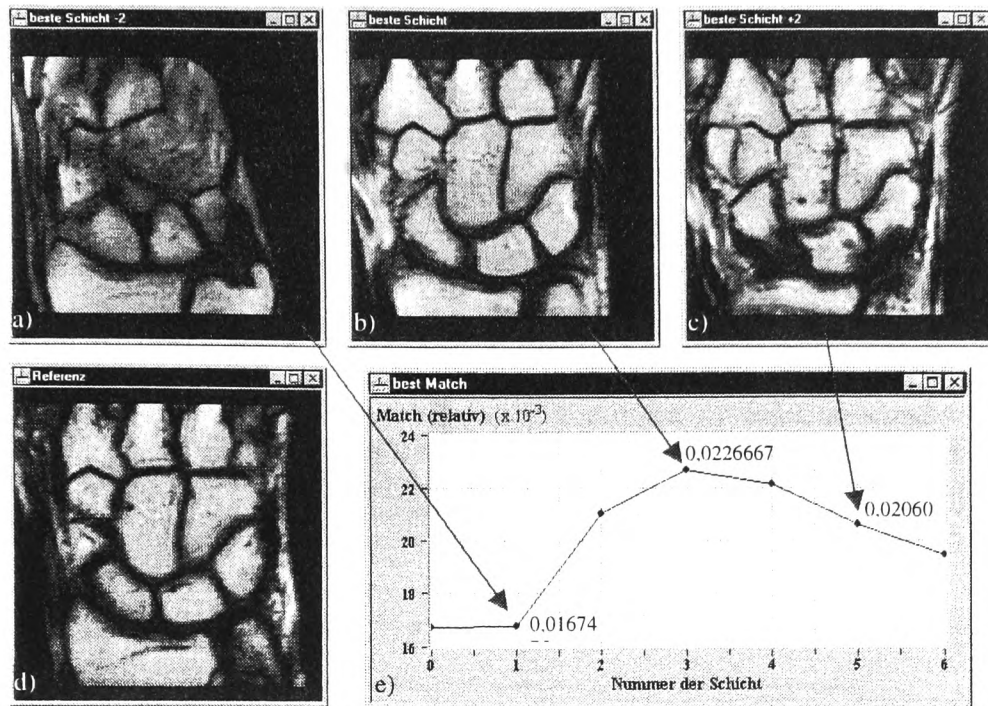
Figure 1: Selection of the best layer: a) upper layer, it does not contain all bones, b) suitable layer, contains intersection of all bones, c) under layer, some bones are partially intersected, d) reference image, e) correspondence of the layers with reference image
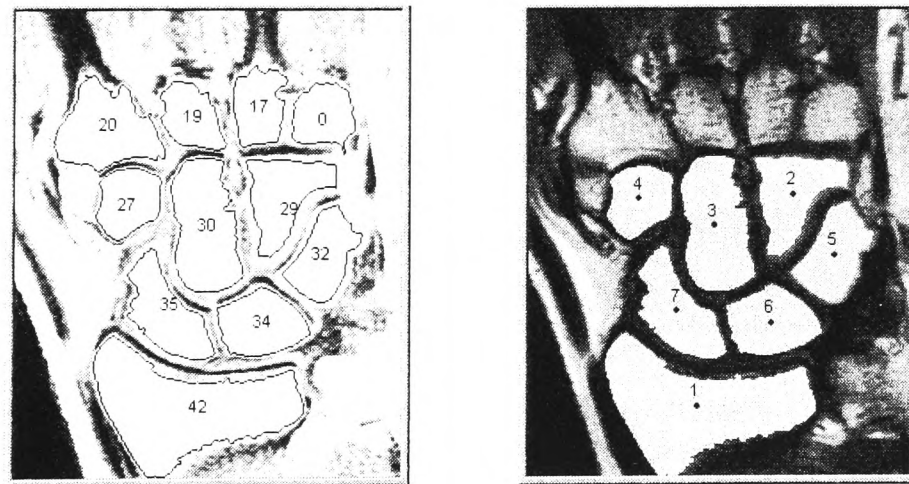

Figure 2: Identification of the wrist bones. a) Pre-selection of candidates through an analysis of the region's position and size. b) Relevant bone regions, automatically identified: (1 Radius, 2 Os hamatum, 3 Os Capitatum, 4 Os Trapezoideum, 5 Os Triquetrum, 6 Os Lunatum, 7 Os Scaphoideum).
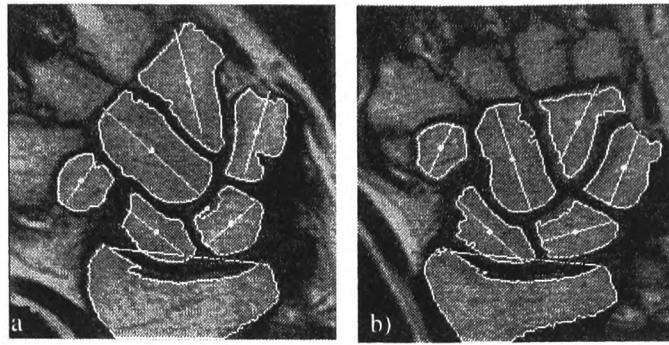
Figure 3: Measurement of translation and rotation - For each bone its major axis and centroid is determined. The anatomic reference co-ordinate system is derived from measuring salient feature points of the Radius. This examples shows three wrist positions: radial deviation, neutral, and ulnar deviation.
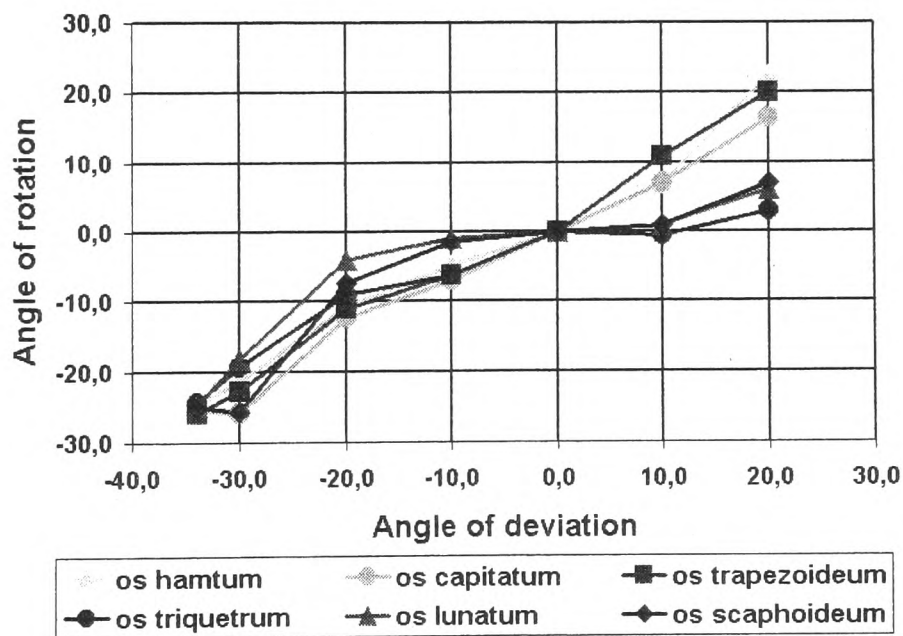

Figure 4: Motion graph – Collection of measurement results for several wrist positions. This graph shows the rotation of the wrist bones for a healthy patient. The rotation of each bone is normalised to its angle at the wrist's neutral position.

References

|1| Larsen, C.F.. "Cineradiography of the wrist", in *Wrist Imaging*, Brunelli, Saffar, pp. 100-106, 1992

|2| Saffar, Ph.. Sokolow, C., Mathoulin, C. Martin-Bouyer, Y.. Verdeille, S., "Cinearthrography of the wrist in carpal instability", in *Wrist Imaging*, Brunelli, Saffar, pp. 109-114, 1992

|3| Linscheid, R.L., Dobyns, J.H., „Karpale Instabilitäten", *Orthopäde*, vol. 22, pp. 72-78, 1993

|4| Peicha, G., Sibert, F.J., Fellinger, M., Grechenig, W., Schippinger, G., "Lesions of the scapholunate ligaments in acute wrist trauma – arthroscopic diagnosis and minimally invasive treatment", *Knee Surgery, Sports Traumatology, Arthroscopy*, vol. 5, pp. 176-183, 1997

[5] Wolf, M., Weierich, P., Niemann, H., "Automatic Segmentation and 3D-Registration of a femoral bone in MR images of the knee", *Pattern Recognition and Analysis*, vol. 7, no. 1, pp. 152-165, 1997

[6] Chen, Q., Defrise, M., Deconinck, F., "Symmetric Phase-Only Matched Filtering of Fourier-Mellin Transforms for Image Registration and Recognition", *IEEE Pattern Analysis ans Machine Intelligence*, vol. 16, no. 12, pp. 1156-1168, 1994

[7] Kobayashi, M., Berger, R.A., Nagy, L., Linscheid, R.L., Uchiyama, S., Ritt, M., An, K., "Normal kinematics of carpal bones: a three-dimensional analysis of carpal motion relative to the radius", *Journal of Biomechanics*, vol. 80, no. 8, pp. 787-793, 1997

FUZZY SNAKES

F Höwing[1,2], D Wermser[2], L S Dooley[1]

[1] University of Glamorgan, United Kingdom
[2] Fachhochschule BS/Wolfenbüttel, Germany

ABSTRACT

A new method for representing and tracking of object boundaries is presented, which allows for the integration of uncertain *a priori* knowledge into an active contour model.

The new fuzzy snake allows for an intuitive specification of the properties of an object's boundary. This is obtained by setting up a linguistic rule base, which describes each of the fuzzy snake's segments. Furthermore the approximate length of each contour segment may be specified to improve the segmentation process and to reduce the computational complexity.

INTRODUCTION

Active contours (Kass et al (1)), or *snakes*, are a well known method for matching an object's contour model to features in an image. The approach, which uses a polygonal object representation (Fig. 1a), is distinguished by its intrinsic ability to handle variations in the boundary to be detected. It is therefore capable of identifying and tracking deformable objects in image sequences.

With the polygonal representation, it is possible to firstly formulate an appropriate energy function E_{snake} for the object. E_{snake} is then minimised to obtain the desired contour by selecting an optimal set of vertices from candidates in a region around an initial contour.

Usually E_{snake} consists of two components E_{int} (internal energy to produce a smooth shape) and E_{ext}. The external energy E_{ext} may be composed of E_{image} which guides the contour towards features in the image and E_{con} which allows for the integration of additional constraints.

Some computational problems resulting from the original algorithm presented in (1) have been overcome by Amini et al (2). The energy minimisation of the snake is performed by a discrete dynamic programming algorithm. This approach allows for the integration of hard constraints such as a minimum distance between the snake's vertices.

An important extension to this work was the introduction of a shape constraint introduced by Höwing et al (3). In addition to only using image features, the integration of geometrical knowledge (within E_{con}) about the object considerably improves detection of structures in images of poor quality (see Wermser and Höwing (4)).
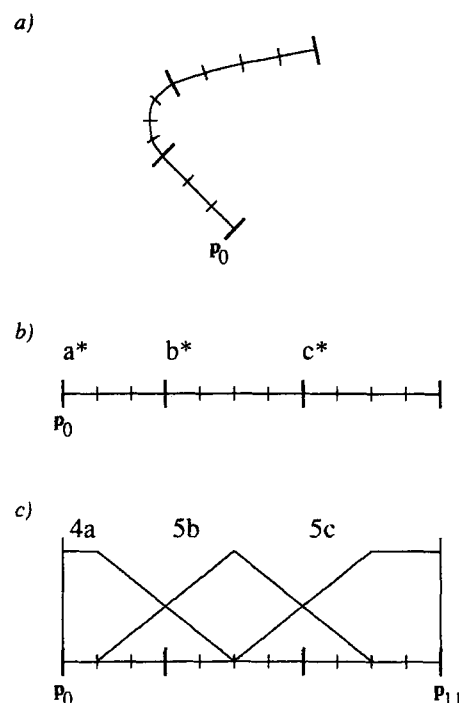

Figure 1: a) Polygonal contour representation. b) Grammatical description of a multi-segment contour. c) Fuzzy specification of the length of contour segments.

ACTIVE CONTOURS WITH MULTIPLE SEGMENTS

There remain some limitations in terms of representing *a priori* knowledge concerning more complex objects which are to be detected. All vertices of the active contour are characterised by the same local energy function, resulting in a single, global description of the object. In order to overcome this problem, Olstad (5) introduced a grammatical description of the snake's energy function. Fig. 1b

shows an example of a contour, which can be described as a sequence of three different (external) energy functions E_{ext}^a, E_{ext}^b and E_{ext}^c, represented by the terminals a, b and c. The grammatical expression describing such a segmented boundary would be a*b*c*, using the closure operator * which allows parts of a pattern to be repeated arbitrarily. A pattern matching algorithm incorporated into the active contour's energy minimisation, constrains the possible resulting contours to comply with the expression.

Three fundamental drawbacks in this algorithm are investigated and subsequently overcome in this paper:

1. The different energy functions do not intrinsically consider inexact *a priori* knowledge.

2. The algorithm is computationally expensive, since the closure operation generates a large number of possible states in the finite-state-machine based pattern matching.

3. The length of a contour segment cannot be specified, although it may approximately be known in advance.

FUZZY SEGMENT LENGTH

The first two problems have a common solution. The fuzzy snake allows for specification of the number of subsequent vertices sharing a common energy function. This dramatically reduces the search space. A crisp length specification however, would not consider uncertain information. A new method to specify the length of a snake segment by a fuzzy number (see Bezdek and Pal (6)) is therefore proposed.

Notation

Instead of specifying snake segments of arbitrary length, for example a*b*c*, the length of each segment is given by the expression aaaabbbbbccccc. Since segment lengths are not known precisely, a different notation is introduced, where a length is given as a fuzzy number, for example $(4a)(5b)(5c)$. Fig. 2 illustrates such a fuzzy number l. In this example the segment specification $(4a)$ denotes a segment consisting of $1 - 7$ vertices, sharing a common property defined by an energy function E_{ext}^a.

Similar to a grammatical description, the fuzzy snake can now be specified as an expression, using operators such as concatenation, logical AND and OR, plus a fuzzy length. The implementation of logical operations will be examined in a forthcoming publication.
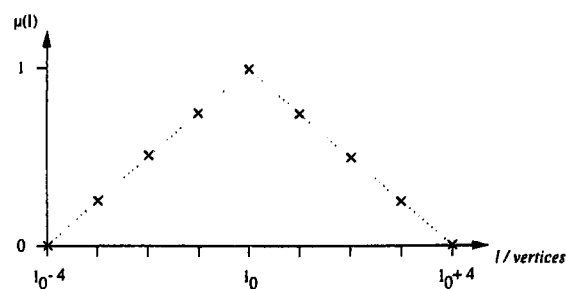

Figure 2: Segment length as a fuzzy number.

Extended algorithm

An extension of the dynamic programming algorithm is necessary to account for the *variable* length of snake segments.

An active contour with only one global energy function is optimised by the (simplified) dynamic programming algorithm in Fig. 3, where N is the number of vertices and M the number of candidates for each vertex.

1. **for** $n = 1 \ldots N - 1$
2.   **for** $m = 0 \ldots M - 1$
3. $S(n,m) = \min_k \big[w_{int} E_{int} + w_{ext} E_{ext}(\mathbf{c}_{n,m})$
   $\qquad\qquad\qquad + S(n-1,k) \big]$
4. $T(n,m) = k^{min}$

Figure 3: Dynamic programming algorithm.

$S(n,m)$ represents the minimal energy level that is possible for the vertices $0, \ldots, n$ if the nth vertex is the candidate $\mathbf{c}_{n,m}$. $T(n,m)$ points to the optimal predecessor of $\mathbf{c}_{n,m}$. After all vertices have been processed, these pointers are traced back, to obtain the new boundary.

TABLE 1 - Energy functions which have to be considered in the example of Fig. 1c

Vertex	0	1	2	3	4	5	6	7	8	9	10	11
Func-	a	a	a	a	a							
tions			b	b	b	b	b	b	b			
						c	c	c	c	c	c	

With multiple external energy functions a different function may have to be calculated for each vertex. Since the segment length is now variable, several energy functions may be calculated simultaneously for vertices on or near a segment boundary. An example is given in Table 1.

Linguistic rules

As described above, each energy function for a boundary segment of constant properties consists of a rule base, which inputs a number of different evidences or features from the image, as well as constraints on geometry and motion of an object. The output of the inference process using this rule base, is a *quality* measure, describing the compliance with the rule base for each vertex.

The contour in Fig. 1a could, for example, be described by three different (simplified) rule bases a, b and c:

a:
IF *curvature* flat
AND *orientation* nw
AND *edge* weak
THEN *quality* very good

b:
IF *curvature* medium right
AND *orientation* north
AND *edge* very weak
THEN *quality* very good

c:
IF *curvature* weak right
AND *orientation* east
AND *edge* strong
THEN *quality* very good

Defuzzification

The defuzzification converts the fuzzy quality measure, which is the result of the inference process, into a crisp value $E^z_{ext}(c_{n,m})$.

CONCLUSIONS

The fuzzy snake is a new form of an active contour, widening the areas of application of this well known contour identification approach. The advantage of the new method in particular, is the ability to make use of uncertain *a priori* knowledge, such as a verbal description of object properties by a human expert. Allowing for a more detailed object description, the fuzzy snake approach presented in this paper improves boundary detection in images of poor quality, such as medical X-ray sequences (4,7). As first results are promising, a forthcoming publication will examine these results as well as details on the implementation.

REFERENCES

1. Kass M., Witkin A., and Terzopoulos D. 1987, "Snakes: Active contour models", Proceedings of the International Conference on Computer Vision, London, 259-268

2. Amini A. A., Tehrani S., and Weymouth T. E. 1988, "Using dynamic programming for minimizing the energy of active contours in the presence of hard constraints", Proceedings of the 2nd International Conference on Computer Vision, 95-99

3. Höwing F., Wermser D., and Dooley L. S. 1996, IEE Electronics Letters, 32, 444-445

4. Wermser D. and Höwing F. 1996, Biomedical Journal, 17-19

5. Olstad B. 1991, "Active contours with grammatical descriptions", Proceedings of 6th International Conference on Image Analysis and Processing, Como, Italy

6. Bezdek J. C. and Pal S. K., eds. 1992, "Fuzzy models for pattern recognition", IEEE Press, New York

7. Höwing F., Dooley L. S., and Wermser D. 1997, "Tracking of non-rigid objects in medical X-ray image sequences", submitted to J. Electronic Imaging