# CONSTRUCTION OF THE SCALE AWARE

# ANISOTROPIC DIFFUSION PYRAMID

# WITH APPLICATION TO

# MULTI-SCALE

# TRACKING

By

C. ANDREW SEGALL

Bachelor of Science
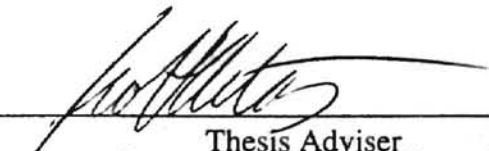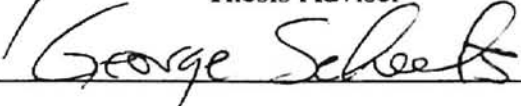
Oklahoma State University

Stillwater, Oklahoma

1995

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December, 1997

# CONSTRUCTION OF THE SCALE AWARE

# ANISOTROPIC DIFFUSION PYRAMID

# WITH APPLICATION TO

# MULTI-SCALE

# TRACKING

Thesis Approved:

_____
Thesis Adviser

George Scheets

K G. Tzaa

Wayne B Powell

Dean of the Graduate College

## ACKNOWLEDGMENTS

Many people have contributed to this thesis, and I would like to specifically recognize several individuals. Dr. Scott T. Acton, my major professor, has supported me throughout this entire project. Though his contributions have been many, I want to thank him for his friendship, constructive criticism, and gentle guidance. Additionally, I want to extend him my appreciation for allowing me, at times, to be self-reliant. This thesis has taught me many things, about academia and about myself, and I attribute the positive result of this learning experience to him.

Other individuals within Oklahoma State University have also affected this document. I want to thank Dr. Keith Teague for his encouragement, friendship, and honesty. I wish to acknowledge Dr. George Scheets for his influence, both in and out of the classroom. I want to express my appreciation to everyone associated with the Oklahoma Imaging Laboratory during my tenure, and I wish to specifically thank Dr. Walter Andrews, Mr. Anthony Wright, and Dr. Wei Chen for their many vibrant discussions. I have enjoyed interacting with so many passionate individuals.

Completion of this thesis has required personal sacrifices, and I want to acknowledge all of my friends and relatives for their love and understanding. Specifically, I want to thank my wife, Sarah, for sharing the peaks and valleys of this

experience. She has helped me attain things I believed unattainable, and this is one of the many reasons that I love her. Thanks also go to my parents, Dr. and Mrs. William E. Segall, for their never-ending encouragement and support.

Finally, I would like to recognize my grandparents, Mr. and Mrs. E.E. Crocker, for their intangible contributions to this document and my life. In fact, I would like to dedicate this thesis to them.

# TABLE OF CONTENTS

# LIST OF FIGURES

# CHAPTER I

## INTRODUCTION

This thesis is concerned with the identification of features within two-dimensional imagery. Current acquisition technology is capable of producing very high-resolution images at large frame rates and generating an enormous amount of raw data. Exceeding present signal processing technology in all but the simplest image processing tasks, the visual information contained in these image sequences is tremendous in both spatial and temporal content. A majority of this detail is relatively unimportant for the identification of an object, however, and the motivations for this thesis, at the core, are the study and development of methods that are capable of identifying image features in a highly robust and efficient manor.

Biological vision systems have developed methods for coping with high-resolution imagery, and these systems serve as a starting point for designing robust and efficient algorithms capable of identifying features within image sequences. By foveating towards a region of interest, biological systems initially search coarse-scale scene representations and exploit this information to efficiently process finer resolution data. This search procedure is facilitated by the nonlinear distribution of visual sensors within a biological vision system, and the result is a very efficient and robust method for identifying objects. Humans will initially identify peripheral objects as potential regions of interest, acquiring higher-resolution image information by focusing on the region, and

1

deciding if the perceived object is actually present through the use of all available knowledge of the scene.

The majority of electronic image sensors, which provide data to signal processing devices, do not utilize nonlinear sensor distributions like their biological counterparts. Attempts to mimic the foveating behavior of a human visual system with software must replicate the advantages of these nonlinear scene descriptions with uniformly sampled data. The first important development in a digital coarse to fine search is encapsulated by continuous *scale space theory*, where it is proposed that an infinite number of scene representations exist and may be created by filtering the original signal with a linear, *scale generating*, filter. The result is a signal description that may be queried in a method similar to the biological coarse to fine search, as objects are initially identified in coarse resolution information, and this information used to guide and refine higher resolution inspections.

Scale space theory provides a fundamental tool in developing object identification search techniques, but it is severely handicapped for actual image processing applications. Construction of scale space is quite expensive, requiring an infinite number of scene representations to follow features from coarse information to finer scene depictions. Additional deficiencies in the application of scale space theory to practical signal processing systems arise in the execution of the initial coarse scale identification, as searching this scene representation is computationally equivalent to searching the original image (which was already deemed excessive). While the introduction of scale space theory is an important first step in designing algorithms capable of replicating biological

vision systems, its direct application to image processing tasks does not reduce the computational requirements of feature identification.

An attractive method for processing scale space data is expressed with an *image pyramid*. Image pyramids are constructed by successively filtering and subsampling the original image, resulting in a data structure that graphically resembles a pyramid. Effectively, these structures contain a discrete representation of scale space, quantizing traditional scale space theory along the scale parameter and quantizing the scale representations along spatial dimensions. Construction of these scene depictions is very efficient, and the initial identification of a feature in coarse scene information comes with reduced computational requirements, compared to searching the original, high-resolution image. Theoretically, the image pyramid will provide a very robust and efficient solution to the object identification problem, allowing an algorithm to query higher resolution imagery only at locations of possible object occurrence and replicating the coarse to fine search method of a biological vision system.

Unfortunately, image pyramid techniques have been largely unsuccessful in actual image processing applications, as traditional scale space theory dictates the use of a Gaussian linear kernel in the generation of the image pyramid. Constructed by successively convolving the original image with a Gaussian-weighted filter and subsampling, the use of the Gaussian pyramid introduces difficulties in the implementation of recognition algorithms. Specifically, linear filtering induces feature movement between spatial representations, yielding poor edge localization, region merging, and increased system complexity. Within a Gaussian pyramid, features of

3

interest may still be identified within a coarse scene representation, but using this information to guide finer resolution inspections has been unfeasible. Gaussian pyramids undersample linear scale space along the scale parameter.

While the practical difficulties inherent to the Gaussian kernel have hindered application of these pyramid structures, their appeal has motivated the relaxation of traditional scale space filter constraints and encouraged the exploration of alternative scale generating mechanisms. Many nonlinear operators have been investigated, but with the necessity of producing a scale space in which regions do not merge and edges do not move, a novel filtering technique has been developed. Modifying an alternative representation of the Gaussian kernel, the anisotropic diffusion smoothing process was designed to allow intra-region smoothing while inhibiting inter-region interactions. The anisotropic diffusion smoothing process attempts to generate scale spaces with minimal feature movement, overcoming the problems inherent to traditional, linear scale space construction.

Development of the anisotropic diffusion equation does address the practical problems for image pyramid application; however, it posses one crucial design flaw. Image pyramid construction (and scale space theory) requires information to be removed as the scale parameter of the filter is increased. The anisotropic diffusion operator, in fulfilling its goal of preserving all candidate edge locations, is unable to guarantee this filter performance. This trait introduces specific problems for the application of the anisotropic diffusion expression to an image pyramid, and it severely compromises the effectiveness of the filtering mechanism in *any* image processing application, as the

nonlinear smoother is unable to remove noise or fine detail. Sensitivity of the anisotropic diffusion equation to noise has received much attention from the research community and several modifications to the diffusion process have been suggested. However, in the development of these modified diffusion algorithms, application of the anisotropic diffusion process to image pyramid representations has been neglected.

The purpose of this thesis is to formalize the construction of an anisotropic diffusion pyramid and apply it to the problem of object identification. This task will require the introduction of, yet, another form of the anisotropic diffusion coefficient and the definition of filter parameters that allow the diffusion process to serve as a suitable prefilter for subsampling operations. Toward this goal, it will be necessary to discuss previous extensions to the diffusion expression, pointing out their respective flaws, and to model the anisotropic diffusion process as a piece-wise linear filter. Construction of an anisotropic diffusion pyramid will also allow this thesis to consider the practical possibilities of replicating biological search methods with digital devices, and will specifically address issues relating to the application of the anisotropic diffusion pyramid to target tracking tasks. The goal of this thesis is to provide the first formal description of an anisotropic diffusion pyramid and apply this new multi-scale structure to the problem of object identification and tracking.

The rest of the paper is organized as follows: Chapter II presents background on multi-resolution image pyramids, scale space theory, and the anisotropic diffusion equation. Chapter III discusses previous extensions to the anisotropic diffusion expression, showing that they are incapable of fulfilling the original goals for the

diffusive process while accomplishing their extended goals. Chapter IV develops a novel diffusion representation, overcoming previous diffusion difficulties, and capable of generating scale space representations with reduced information content and with minimal feature drift. Chapter V formalizes the construction of the anisotropic diffusion pyramid, developing conditions for all filter parameters that allow the anisotropic diffusion process to serve as a suitable sampling prefilter. Chapter VI applies the new pyramid structure to an object identification and tracking system, presenting numerical simulations which display the performance possibilities of the anisotropic diffusion pyramid in a target tracking application. Chapter VII presents conclusions and future work.

# CHAPTER II

# BACKGROUND

## Overview

Edges are an integral part of most image processing applications. Through the recognition and location of these features, raw visual information is transformed into semantically meaningful representations and expressed in compact forms. This collection of edge information allows a vision system to query the data, identify objects, and determine their locations; all are fundamental tasks of an image processing system.

Computational approaches towards locating edges are varied, but common methods incorporate first or second derivatives in the procedure and locate either local extrema or zero crossings. While these operations do provide meaningful measurement of region boundary existence, their application is problematic, as differentiation is sensitive to fine features, texture, and noise. Traditionally these spurious responses are suppressed by prefiltering the signal before differentiation, equivalent to enlarging the *scale* of the derivative calculation.

Introducing a prefilter into the detection operation presents problems for the theoretical development of edge identification. Specifically, one would like to determine if there exists an optimal prefilter for the detection procedure and, if so, what is its ideal scale? Pursuit of these answers forms the background of this thesis, and the goal of this

section is to provide an introduction to several important concepts. The section begins with discussion of scale space theory and continues by presenting image pyramids as its discrete extension. Problems applying continuous scale space concepts within image pyramids necessitates work on novel filtering techniques, and the section concludes with the introduction of one such adaptive filtering mechanism, the anisotropic diffusion equation.

Scale Space

Scale space filtering was initially developed to manage the relationship between edge information over varying resolution. Since many signal characteristics, most notably derivatives, are calculated over a region where the region size influences the descriptive measurement, Witkin introduced scale space as a collection of signal representations, derived from the original image and generated by a scale space filter [37]. Scale space does not attempt to define an optimal scale for feature identification, but provides a method for establishing correspondence between edges found in heavily filtered signal representations and their location in the original signal.

Construction of scale space is straightforward, and traditionally begins by filtering the original signal with an FIR filter of varying width. Hyper-planes within scale space contain a single filtered representation of the signal, while filtering the signal with a continuum of filter widths produces scale space. For a two-dimensional image, scale space may be visualized as a three-dimensional cube, containing an infinite number of

signal descriptions stacked upon one another. These representations are ordered by their

respective filter scale parameter, and from these filtered images, features may be

identified using traditional detection methods. An example scale space is shown in

Figure 2.1.
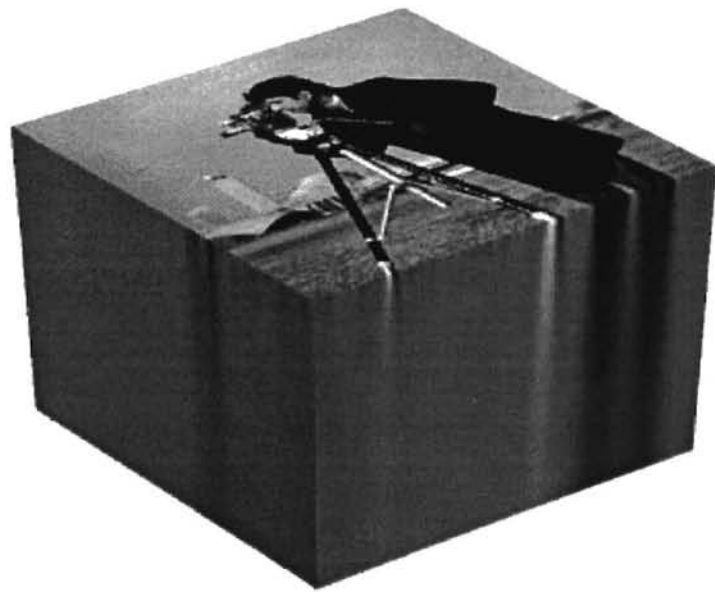


**Figure 2.1** Scale space for the cameraman image. The original image is located at the top of the cube and lower levels are occupied by coarser representations of the scene. For this example, the coarse scale images are constructed by smoothing the original image with a Gaussian filter.

Plotting the locations of detected features versus the continuous scale parameter is defined as the *scale space image* and an example is displayed in Figure 2.2. Within these structures, objects may be recognized at coarse resolution representations and traced to their origin. This is referred to as a *coarse to fine search* and encompasses the power of scale space theory, allowing the initial identification of significant features to occur in the absence of spurious derivative results. The exact location of these edge points in the original image may then be obtained by traversing scale space towards finer resolution, resulting in a robust method for fusing multi-scale information and a procedure well suited to the edge detection problem.
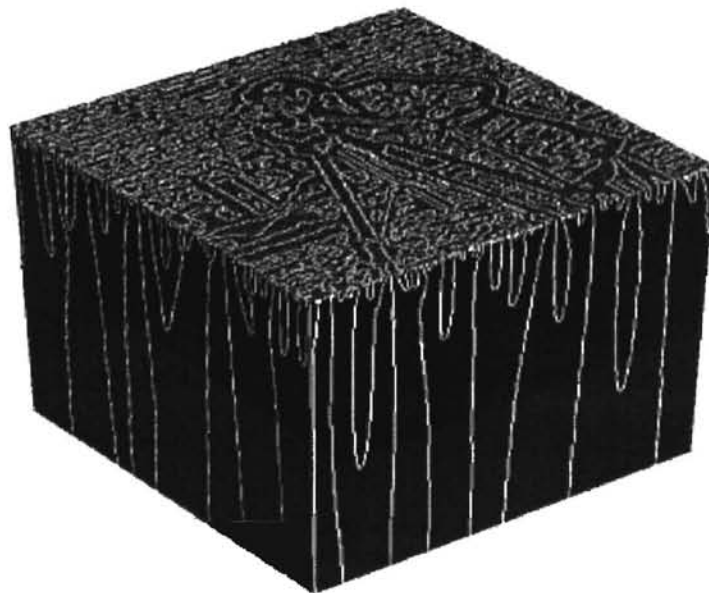


**Figure 2.2** Scale space and the scale space image. Scale space images are displayed on the sides of the cube and show how smaller features disappear rapidly as scale is increased. Connectivity between levels is defined as spatial causality, as all coarse scale features correspond to features in the higher resolution representations.

Specification of a viable scale space filter requires fulfilling a specific smoothing criteria: if a feature is tracked across increasing scale, it should disappear. Conversely, a new feature should never appear while scale increases, as coarse resolution representations would no longer correspond to the original signal. Guaranteeing the presence of coarse scale objects in finer scene representations is expressed as *spatial causality*, maintaining a cause and effect relationship for features, and is a necessary condition for application of the multi-scale coarse to fine search method.

The spatial causality criterion allows the specification of an optimal filter for scale space generation. Witkin initially restricted the scale generating filter to be symmetric, strictly decreasing about the mean, and linear. As a result of this definition, it has been shown that the Gaussian kernel is the only filter capable of satisfying these constraints in one-dimension while maintaining spatial causality [3]. The uniqueness of the Gaussian kernel for scale space construction has since been extended to higher dimensions [41], discrete signals [23], and the larger class of unsmooth kernels [39]. The Gaussian filter also has the unique property of minimizing the uncertainty principle [25].

Even with the discovery of a unique scale generating filter, application of scale space theory to practical problems is limited. Requiring an infinite (or near infinite) number of scale representations necessitates large storage requirements, and performing feature detection tasks on each resolution level is computationally expensive. Efficient execution of a coarse to fine search demands quantization of the scale parameter, and is formalized by the image pyramid.

Image Pyramids

Image pyramids are a discrete representation of scale space. By requiring the calculation of fewer scene representations, they reduce the computational requirements of scale space construction and the coarse to fine search. Image pyramids also introduce additional processing speed-up by coupling their choice of scale retention to the sampling properties of a scale generating operator. Allowing the decimation of coarse resolution representations results in decreased storage requirements, faster scale space construction, and a logarithmic improvement in coarse to fine matches. Theoretically, an image pyramid will provide a very efficient *and* robust solution to the feature identification problem.

Construction of an image pyramid begins by filtering the original signal. This coarser resolution representation, now satisfying some sampling criterion, is then decimated, which traditionally consists of discarding all pixels belonging to the even rows and columns of the image. Subsequent pyramid levels are created by iteratively filtering and subsampling the previous resolution representation, and the product is a set of image descriptions, each of smaller size than the original. Sorting these images according to scale is shown graphically in Figure 2.3 and presents the visual appearance of a pyramid structure, hence the name image pyramid.
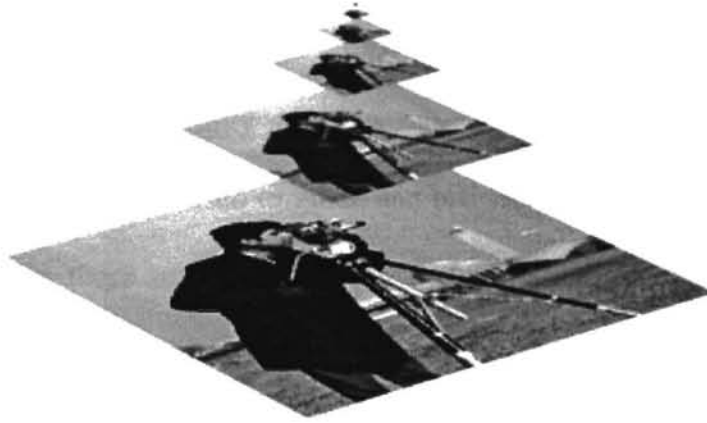
**Figure 2.3** An image pyramid, constructed by filtering the original cameraman image and subsampling. The original image is located at the bottom of the pyramid and coarser scale representations occupy successively higher pyramid levels. For this example, the coarse scale images are constructed by smoothing the original image with a Gaussian filter.

Gaussian pyramids correspond to the Gaussian generated scale spaces of the previous section, and they are constructed by using a Gaussian kernel as the scale generating filter and applying Shannon's sampling theorem for the decimation operation. Mathematically, the construction of pyramid level $L$ of image $I$ can be described as

$$\mathbf{I}_L = \left[\mathbf{G}_\sigma * \mathbf{I}_{L-1}\right]_{\downarrow S}, \tag{2.1}$$

where $\mathbf{G}_\sigma$ is a Gaussian scale generating filter of standard deviation $\sigma$, $\downarrow S$ denotes subsampling by a factor of $S$ within each row and column, and $\mathbf{I}_0$ is the original image. To remove frequencies below the Nyquist rate, the Gaussian filter is defined to have width $2S/\pi^\dagger$ [5].

---

$^\dagger$ In this derivation, the high frequency limit of the Gaussian kernel is defined to be $(\pi\sigma)^{-1}$, equivalent to twice the standard deviation of the filter in the frequency domain.

Generation of an image pyramid provides significant performance enhancement for the coarse to fine search. Initialized on the coarse resolutions in the pyramid, the search procedure needs to only fully search a subsampled image representation. The result of this first search is then used to guide and refine progressively higher resolution inspections, restricting these subsequent examinations to small regions within the next scale description. Ignoring the effort for pyramid construction, the computational increase of searching a multi-scale structure, compared to inspecting the original full resolution image, has been shown to be $S^{2L}$, where $S$ is the sample factor and $L$ is the coarsest level of the search [38]. The performance improvement between an image pyramid and traditional scale space would be the same, actually amplified by the decreased construction costs of an image pyramid.

With the promise of computational efficiency and the inherent robustness of scale space, use of multi-scale image pyramids has been widespread. Burt has discussed their application to motion analysis, object recognition, and selective transmission [7]. They have also been applied to image compression [8], scene matching [38], target identification [38], segmentation [10], multi-resolution splines [9], and texture synthesis [27]. Reducing computational requirements by orders of magnitude, real time implementations using these pyramid structures have been presented with application to feature identification [6,11].

Unfortunately, the wide application of pyramid structures has shown that the Gaussian kernel, previously defined as optimal, possesses several undesirable characteristics for quantizing the theory of scale space within a computer vision system.

14

Specifically, as the scale parameter of the filter is increased, regions tend to merge and edges move due to the lowpass response of the filter. In continuous scale space, movement between resolution representations presents no obstacle to the coarse to fine search and is accommodated by allowing minimal scale change between neighboring levels. Image pyramids contain a small number of scale representations, reducing the number of scales available to a pyramid search procedure, and with fewer scale depictions, large feature movement between pyramid levels makes attempts at following a feature across scales ineffective. This dramatically decreases system robustness and performance.

Robust and efficient solutions to the feature identification problem have application to a large class of image processing tasks, and the theoretical promise of an image pyramid to deliver this quality of solution encourages the relaxation of initial restrictions on the scale generating kernel. Exploring the use of alternative scale generating filters, pyramids have been constructed with Haar wavelets [12] and morphological operators [33], both with the goal of generating scale spaces immune to feature movement while ensuring spatial causality. A filtering process designed specifically for this purpose is the anisotropic diffusion technique, and this technique will be introduced in the next section.

# Anisotropic Diffusion

Evolution of the anisotropic diffusion equation originated as an enhancement of traditional linear filtering. The class of linear filters is widely used in signal processing and theoretically well developed; however, its characteristic global smoothing blurs region boundaries and results in edge removal or feature drift, undesirable attributes for tasks such as image pyramid construction. The anisotropic diffusion equation modifies the behavior of one member of the linear filter class, the Gaussian kernel, and creates a process which adaptively smoothes within regions while inhibiting intra-region interaction.

In the traditional scale space representation, a Gaussian pyramid is usually constructed by convolving the original image with a suitable Gaussian kernel and subsampling. This multi-scale structure can also be implemented with the use of the heat equation [22]. For an image defined on a discrete grid, this process is approximated by the partial differential equation

$$\frac{\partial \mathbf{I}}{\partial t} = \nabla^2 \mathbf{I},$$

(2.2)

where $\mathbf{I}$ is the input signal, $\nabla^2$ is the discrete Laplacian, and $t$ is the solution time or scale parameter.

Solution of the heat equation is completely defined by its Green's function, the Gaussian kernel, with the width of the resulting kernel proportional to solution time [36]. In creating a scale generating process with the capability of maintaining region integrity,

16

the heat equation may incorporate a spatially varying damping coefficient. The modified process becomes

$$\frac{\partial \mathbf{I}}{\partial t} = div(\mathbf{c} \cdot \nabla \mathbf{I}), \tag{2.3}$$

where $\mathbf{I}$ is the input signal, $div$ is the divergence operator, $\nabla$ is the discrete gradient, and $\mathbf{c}$ is the adaptive diffusion coefficient. For a two dimensional image, one possible realization is

$$\Delta \mathbf{I} = \Delta t \cdot \left( c_N \nabla_N \mathbf{I} + c_S \nabla_S \mathbf{I} + c_E \nabla_E \mathbf{I} + c_W \nabla_W \mathbf{I} \right), \tag{2.4}$$

where $\Delta t$ is the discrete solution time step, $\nabla_N$, $\nabla_S$, $\nabla_E$, and $\nabla_W$ are the gradients (simple differences) in the north, south, east, and west directions, respectively, and $c_N$, $c_S$, $c_E$, and $c_W$ are the diffusion coefficients in the north, south, east, and west directions, respectively [29]. These coefficients are traditionally bounded to the set [0,1] and decrease with increasing gradient, effectively inhibiting smoothing in regions of possible edge location. An example of filtering an image with the anisotropic diffusion equation is presented in Figure 2.4.

**Figure 2.4** A visual example of filtering with the anisotropic diffusion equation. The original image is located at the top, and its filtered result is located below. Notice how the diffusion process smoothes within the boundaries of an object but preserves the edge locations. Maintaining all features with high contrast, especially the smaller objects near the road, will motivate the exploration of spatially aware variants of the anisotropic diffusion process.

Construction of the anisotropic diffusion coefficient defines the performance of the nonlinear filter. With the initial goal of preserving regions of possible transition, varying the coefficient relative to local gradient is well motivated, and in the introduction of anisotropic diffusion as a filtering process, two possible realizations of the diffusion coefficient are suggested [29]. The first is of Gaussian shape and expressed as

$$c(\nabla I) = e^{-\left(\frac{\|\nabla I\|}{k}\right)^2},$$ 
(2.5)

while the second suggestion is

$$c(\nabla I) = \frac{1}{1 + \dfrac{\|\nabla I\|}{k}}.$$
(2.6)

In both diffusion coefficients, a gradient threshold, $k$, is introduced and its selection quantifies the minimum gradient magnitude which should be preserved by the smoothing mechanism.

Alternative coefficient realizations have been derived by studying the stability of the diffusion difference equation. Researchers have found the initial diffusion coefficients in (2.5) and (2.6) to produce a system which is ill-posed, in the sense that small changes in the original image may result in large changes to the filtered result. Conditions guaranteeing that the diffusion process is not ill-posed have been presented by You *et al.* [40], who show that evaluating the derivative of the coefficient at infinity will determine the suitability of a diffusion expression as a scale generating operator. One potential coefficient implementation, satisfying these criteria, is

$$c(\nabla \mathbf{I}) = \begin{cases} \dfrac{p(k+\varepsilon)^{p-1}}{k}, & \|\nabla \mathbf{I}\| < k \\ \dfrac{p(\|\nabla \mathbf{I}\|+\varepsilon)^{p-1}}{\|\nabla \mathbf{I}\|}, & \|\nabla \mathbf{I}\| \geq k \end{cases} \quad (2.7)$$

where $p$ is suggested to be 0.5, $\varepsilon$ is suggested to be 1, and $k$ is the gradient threshold.

Besides research of diffusion coefficient construction, theoretical developments relating to other diffusion parameters have also been presented. Work on the proper selection of gradient thresholds [35, 1] and solution stopping time criteria have been discussed [14, 4]. Additionally, concerns over the computational complexity of an iterative scale generating filter have motivated studies of more efficient implementations of the anisotropic diffusion process, including analog realizations [30], neural networks [15], and mathematical multigrid methods [32, 2].

Anisotropic Diffusion and Scale Space

Application of this new anisotropic diffusion process does produce a scale space void of edge movement, as depicted in Figure 2.5, and early applications include edge detection, scene segmentation, and stereo matching [32]. Coarse to fine search procedures in this new scale space no longer require the computation and storage of large numbers of scale representations, unlike traditional Gaussian generated scale space. Instead, only coarse information must be computed and examined, as features found in these scene descriptions will correspond to features occupying identical spatial locations in the original image. With limited feature movement, it appears that the anisotropic

20

diffusion equation produces a scale space well suited for scale quantization, as many

scales must be traversed before a feature deviates from its original location.
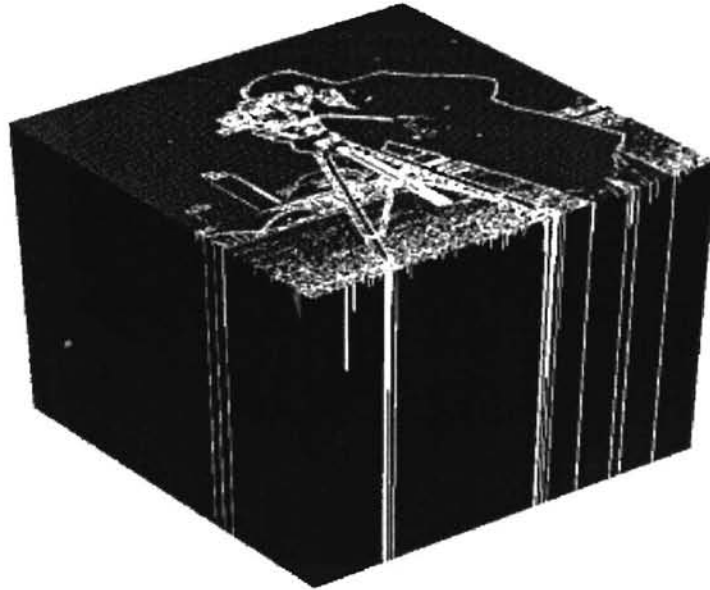


**Figure 2.5** Anisotropic diffusion and its scale space image. The anisotropic diffusion process attempts to smooth while maintaining edge locality. Unlike the traditional linear scale spaces, features within these scale space images do not move across increasing scale, easing the problem of determining feature correspondence.

In creating an image pyramid using anisotropic diffusion, one would successively

diffuse and then subsample the original image. Unfortunately, the filtering performance

of the diffusion mechanism does not satisfy requirements for image pyramid construction,

since signals filtered with the nonlinear anisotropic diffusion equation no longer display a

systematic removal of information and cannot satisfy traditional sampling theorems.

Without the assistance of an image pyramid, creating coarse scene information with the anisotropic diffusion mechanism is computationally expensive and searching these scale representations provides no performance increase, as the resolution of the original and coarse scene information is identical. At best, scale spaces constructed with traditional anisotropic diffusion present a robust solution to the edge identification problem, but at *increased* computational cost.

More troubling for the application of anisotropic diffusion to signal processing tasks is that the diffusive process does not guarantee the removal of *any* information. For example, consider the one dimensional pulse train shown in Figure 2.6, where the pulse heights are identical and defined to be greater than the gradient threshold, $k$, of the diffusion coefficient. Anisotropic diffusion is designed to preserve regions of high gradient, and the traditional definitions for damping coefficients, presented above, result in coefficient values near zero at the pulse edges. Since the diffusion update depends on a weighted sum of the product of local coefficient and gradient magnitudes, smoothing will not occur, as each
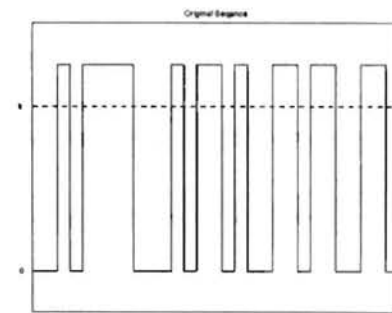


(a)
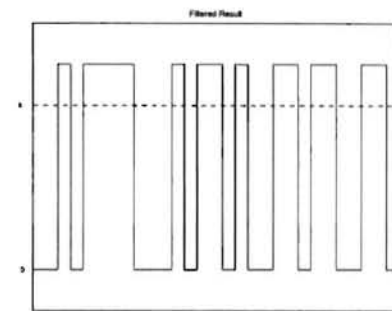
(b)

**Figure 2.6** An input sequence which will not be smoothed by the diffusion process: (a) the original sequence and (b) its filtered result. The gradient threshold of the diffusion system is represented by $k$, and all gradients larger than this threshold are defined to be preserved.

22

signal location posses a gradient or coefficient value of zero. In this example, the filtered result will be identical to the original, and changes in solution time, or scale, do not affect information removal.

Practical filtering problems are motivated by the need to remove noise and spurious detail, as encompassed with scale space theory, and it is apparent that traditional anisotropic diffusion expressions do not accomplish this, instead preserving all regions with high contrast. This characteristic introduces difficulties for the application of anisotropic diffusion to any filtering problem, and researchers have spent time addressing it. The result is a modification to the diffusion coefficient, creating a diffusion equation which is *spatially aware*. In these modified diffusion expressions, the goals of the nonlinear smoothing process are expanded, seeking to preserve features of high gradient *and* to remove regions of small spatial scale.

Spatially Aware Anisotropic Diffusion

Scale of a diffusion equation is enlarged by increasing the scope of the diffusion coefficient calculation. By incorporating greater spatial knowledge of the signal into the decision to diffuse, the filtering process is allowed to smooth small regions regardless of local contrast. A method providing the anisotropic diffusion equation with a direct specification of scale was first proposed by Catté *et al.* [13], who suggest utilizing a Gaussian kernel to spatially expand the coefficient computation. Using the original coefficient expression in (2.5) as an example, a possible spatially aware diffusion coefficient is specified as

$$c(\nabla \mathbf{I}) = e^{-\left(\frac{\|\mathbf{G}_\sigma * \nabla \mathbf{I}\|}{k}\right)^2}, \qquad (2.8)$$

where $\mathbf{G}_\sigma$ is a Gaussian kernel with standard deviation $\sigma$.

Other linear filters have been proposed to replace the Gaussian kernel in (2.8) [34], and proper selection of a spatially aware anisotropic diffusion coefficient is usually motivated by underlying assumptions of the noise distribution within the original signal. While the use of a linear filter within the diffusion coefficient may be viewed as "against the spirit of anisotropic diffusion" [40, 21], initial filtered results using these spatially aware expressions display their ability to remove small regions of high contrast. Figure 2.7 presents a visual example of the smoothing performance of a scale aware anisotropic diffusion process, showing that filters using the spatially enlarged diffusion coefficients are capable of smoothing, and eventually removing, noise.

Development of a spatially aware diffusion equation provides the first step in defining a pyramid structure based on an anisotropic diffusion process, since necessary information removal will be realized by enlarging the scale of the coefficient calculation and increasing the solution time of the diffusion equation. Admittedly, enlarged coefficient expressions will allow features slight movement; however, these coefficients present information reduction which has a more natural meaning of scale and is possibly suitable for sampling. If small objects exist throughout scale space, subsampling coarser resolution representations will result in severe aliasing and compromise the robustness of a coarse to fine search. Spatially aware diffusion variants remove these objects and their inherent difficulties.

**Figure 2.7** A visual example of filtering with the spatially aware anisotropic diffusion equation. The original image is located at the top, and its filtered result is located below. This diffusion process is capable of smoothing regions of small spatial size and high contrast, as evident by the removal of the small objects near the road. Spatial smoothing does reintroduce interaction across region boundaries, and this will be discussed in the next chapter.

A major obstacle in constructing image pyramids with spatially aware anisotropic diffusion expressions is that these new smoothing operators posses multiple definitions of scale. Traditional diffusion equations contain a single scale parameter, corresponding to solution time. Spatially aware anisotropic diffusion operations incorporate a second scale parameter, describing the region over which a diffusion coefficient is calculated. This thesis will discuss image pyramid construction within the context of sampling and spatial causality, and before it is possible to define filter criteria allowing the anisotropic diffusion equation to smooth a signal and produce a result suitable for sampling, it will first be necessary to model the smoothing properties of the scale aware anisotropic diffusion expression.

In the next chapter, analysis of the scale aware diffusion processes, which incorporate linear filters into the diffusion coefficient, will be presented. The goal of this chapter will be to show that these coefficient expressions are incapable of removing regions of small size while preserving edges and possess smoothing characteristics which are ill-suited for pyramid construction. Subsequent chapters will abandon the scale aware diffusion coefficients presented in the literature and will develop a new coefficient expression using nonlinear morphological filters. These chapters will show that the new diffusion operations are capable of simultaneously smoothing a signal while maintaining edge locations and, therefore, suitable for image pyramid construction.

# CHAPTER III

# SPATIALLY AWARE ANISOTROPIC DIFFUSION

## Overview

In the previous chapter, anisotropic diffusion was established as a smoothing filter, capable of reducing the information content of a signal while preserving significant discontinuities. Filtered representations generated with this anisotropic diffusion process allowed for the trivial identification of region boundaries, but without inducing feature movement. This property made the anisotropic diffusion equation applicable to many image processing tasks and provided the appealing possibility of encapsulating scale space theory within an image pyramid.

Unfortunately anisotropic diffusion does not guarantee the removal of noise or fine detail, as original implementations of the diffusive mechanism require the computation of signal characteristics in a limited neighborhood. This initially makes the anisotropic diffusion equation unsuitable for image pyramid construction, but methods have been suggested which increase the scale of the diffusion operation and allow it to smooth small features. In these modifications, the filtering process employs spatially enlarged gradient approximations within its diffusion coefficient, calculated from weighted averages of the underlying signal.

The aim of this thesis is to formalize the construction of an anisotropic diffusion pyramid and apply it to the problem of target identification and tracking. Toward this goal, the purpose of this chapter is to develop a theoretical understanding of the smoothing performance of these new spatially aware anisotropic diffusion processes, since the construction of an image pyramid requires a diffusion operation which is capable of preserving edges while allowing the removal of information.

The goal of this chapter is to show that the use of any linear filter within the diffusion coefficient creates a process *incapable* of simultaneously removing small scale regions while preserving edges and, therefore, *inapplicable* to the construction of image pyramids. This will be accomplished by studying the smoothing of several prototype edge functions and, in the larger structure of this thesis, will motivate the introduction and analysis of diffusion coefficients which incorporate nonlinear filters. These new coefficient variants will be discussed in the following chapter.

Preceding study of the scale aware anisotropic diffusion equation, this chapter will present several analytical tools which simplify its analysis. Specifically, discussion will be presented on the separability of the anisotropic diffusion equation, reducing the two-dimensional equation into a sequence of one-dimensional operations. Additionally, invariance of the smoothing process with respect to coefficient design will be developed through study of the traditional anisotropic diffusion operation, allowing a generalized model for all diffusion systems to be used during theoretical discussion.

This chapter is organized as follows: The first section introduces the two prototype edge functions used throughout the discussion. It also considers the

separability of the anisotropic diffusion equation and provides an overview of the spatially aware diffusion process. The second section abstracts analysis of diffusion realizations from the design and construction of their diffusion coefficient. This is realized through study of the traditional anisotropic diffusion equation. The final section of this chapter explores the suitability of incorporating linear filters into the diffusion equation and their affect on filtering performance. This section will show that utilizing linear filters in the diffusion equation *does not* result in a diffusion process capable of removing features of small scale and low contrast, realizing the goal of this chapter.

## Background

Before exploring the smoothing performance of the spatially aware anisotropic diffusion equation, it is first convenient to concisely present several definitions which will be used in the following discussion. This section consists of three distinct topics. It begins by defining two basic edge models, introducing both their spatial and gradient formulations. The second part of this section briefly derives a separable representation of the anisotropic diffusion equation and introduces an alternative gradient based realization of the diffusion operator. The third section outlines the computation of traditional and spatially aware anisotropic diffusion, developing terminology used in the study of linear filters incorporated into the scale aware coefficient expressions. These subsections present fundamental tools, to be utilized in the subsequent sections.

29

## Edge Function Definition

The two prototype edge models considered in this chapter are the step and finite width step functions. The step function is defined in one dimension as

$$u(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}, \qquad (3.1)$$

and its equivalent discrete gradient representation is calculated using a right-hand difference approximation, expressed as

$$\nabla u(x) = \begin{cases} 1 & x = -1 \\ 0 & otherwise \end{cases} \qquad x \in Z. \qquad (3.2)$$

The edge location of this model is defined to be where the gradient is non-zero, which for the discrete step function is located at $x = -1$. Figure 3.1 graphically summarizes the characteristic properties of the step edge model.
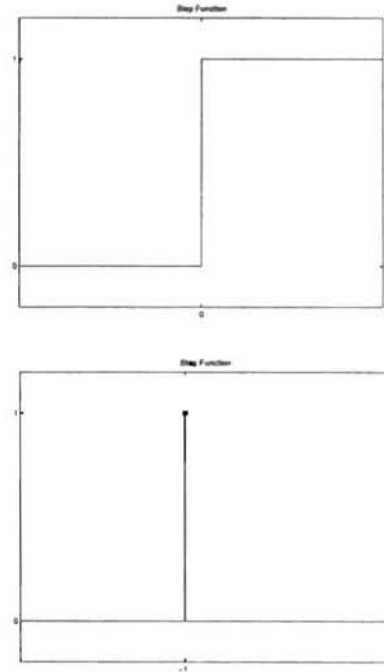


**Figure 3.1** The step function and its equivalent gradient representation.

Combining two step edges produces the second edge model, the finite width step function. This edge function is defined as

$$u_{fw}(x) = u(0) - u(W), \qquad (3.3)$$

where $W$ is the width of the step, and its gradient representation is expressed as

$$\nabla u_{fw}(x) = \delta(-1) - \delta(W-1), \qquad (3.4)$$

30

where $\delta$ is the delta function. The characteristic properties of the finite width step function are summarized graphically in Figure 3.2.



## Separability

Presentation of the two edge models in the previous subsection occurred completely in one dimension, and though this thesis is concerned with the filtering of two-dimensional images, this is acceptable since, like the

**Figure 3.2** The finite width step function and its equivalent gradient representation.

original Gaussian kernel, one iteration of the discrete anisotropic diffusion equation may be viewed as a separable filter. This allows a two-dimensional process to be analyzed as sequential one-dimensional operations and simplifies both theoretical and computational investigation.

Implementing the diffusion update with horizontal and vertical components produces the system:

$$\mathbf{I}_{t+\Delta t} = \mathbf{I}_{t+\Delta t_H} + \Delta t \cdot \left( c_N \nabla_N \mathbf{I} + c_S \nabla_S \mathbf{I} \right), \qquad (3.5)$$

and

$$\mathbf{I}_{t+\Delta t_H} = \mathbf{I}_t + \Delta t \cdot \left( c_E \nabla_E \mathbf{I} + c_W \nabla_W \mathbf{I} \right), \qquad (3.6)$$

where $\mathbf{I}_{t+\Delta t_H}$ is the filtered result after smoothing in the horizontal direction.

31

Separability leads to another useful representation of the anisotropic diffusion process and allows its expression completely in terms of gradients. This eases visualization and further facilitates analysis of the filter goals, smoothing while preserving locations of significant discontinuity. In the one-dimensional diffusion equation (3.6), $\nabla_W I$ may be defined as $\nabla I[x]$. Realizing that $\nabla_E I$ is equal to $-\nabla I[x-1]$, the change in gradient between diffusion updates is expressed as

$$\nabla I_{t+\Delta t}[x] = \nabla I_t[x] + \Delta t \cdot (c[x-1] \cdot \nabla I[x-1] - 2c[x] \cdot \nabla I[x] + c[x+1] \cdot \nabla I[x+1]), \quad (3.7)$$

where $\nabla I[x]$ is the discrete right-hand derivative approximation at location $x$ and $c[x]$ is the spatially adaptive diffusion coefficient calculated at location $x$ and unbiased with respect to gradient direction. A similar representation may be found for the vertical component of the diffusion equation (3.5), replacing the spatial variable $x$ with $y$. These expressions are formally derived and presented in Appendix A.

Linear Filters

Diffusion coefficient expressions are traditionally defined to depend solely on local gradient information, with damping coefficient functions of the form $c(\nabla I)$. Inspection of the diffusion equation presented in (3.7) reveals a departure from original diffusion coefficient expressions, as the anisotropic diffusion coefficient is defined relative to the spatial variable $x$, and this facilitates the use of non-traditional diffusion coefficient realizations. The focus of this entire chapter is the study of diffusive processes that incorporate linear filter into their diffusion coefficients, and preliminary to this analysis, this subsection will provide an illustrative guide to the implementation of

the scale aware diffusion process. In doing so, initial restrictions on the candidate linear filters will be clarified and the nomenclature that describes the linear filter characteristics will be defined. These definitions will be used consistently throughout this thesis.

Original anisotropic diffusion expressions attempt to preserve regions of high contrast and smooth regions with low gradient magnitude. A graphical outline of a prototypical diffusion process is presented in Figure 3.3, and begins by computing the difference signal of the unfiltered sequence. (The example sequence and its difference signal are shown in Figure 3.3a and 3.3b, respectively.) The next task within a diffusion process is to compute the diffusion damping coefficient, and in a traditional diffusion expression, coefficient values are dependent solely on the initial difference signal. Low values of the coefficient allow locations to be preserved and, referring to the example
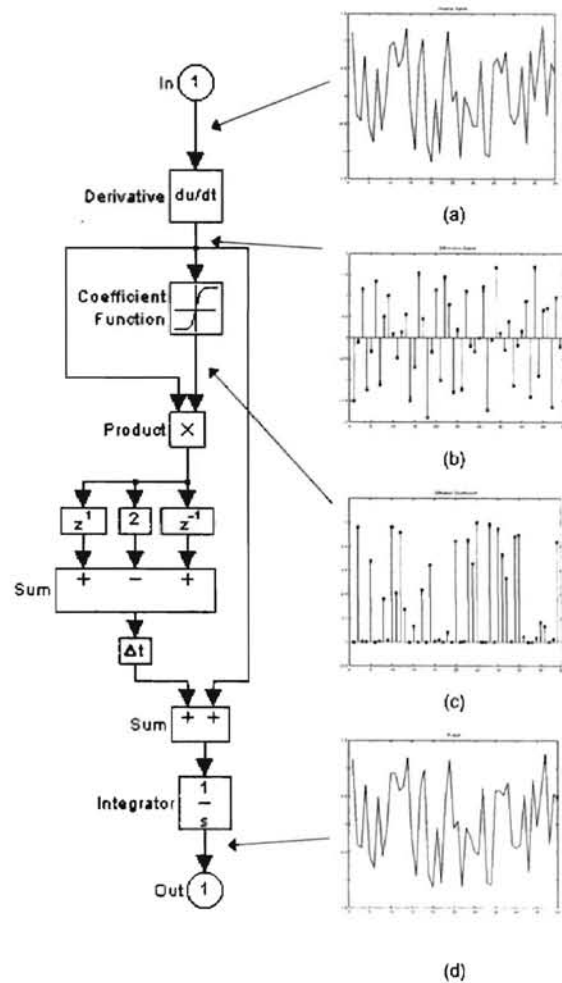


**Figure 3.3** A system representation for one iteration of the traditional anisotropic diffusion process: (a) the input signal, (b) its difference representation, (c) diffusion coefficients, and (d) the filtered result.

33

coefficient result presented in Figure 3.3c, correspond to candidate edge features. Multiplying individual gradient values by their corresponding diffusion coefficient completes the diffusion process, and using the update equation presented in (3.7), the result of diffusing the example sequence with one iteration of a traditional anisotropic diffusion realization is displayed in Figure 3.3d.

Problems arise when attempting to smooth impulsive noise with the traditional anisotropic diffusion process, as these regions contain objects of high contrast that are maintained by the diffusive mechanism. Spatially aware anisotropic diffusion attempts to overcome this deficiency by incorporating a linear filter into the diffusion coefficient, and this effectively increases the spatial scale over which the damping values are calculated. Performing one iteration of the spatially aware anisotropic diffusion process is displayed in Figure 3.4 and is initially similar to traditional diffusion methods. The process begins with the calculation of difference signals from the initial input sequence, and these signal representations are shown in Figures 3.4a and 3.4b, respectively. The next step in the diffusive mechanism is the computation of the diffusion coefficient, and it is in the coefficient calculation that the diffusion processes differ. In a spatially aware anisotropic diffusion process, gradients used within the diffusion coefficient are first filtered with a linear kernel, while in traditional diffusion expressions, they are left unmodified. The result of a possible coefficient prefilter is shown in Figure 3.4c, and the coefficient calculation is shown in Figure 3.4d. Unlike traditional anisotropic diffusion coefficients, this example coefficient sequence no longer produces low smoothing values at all

locations of high contrast. Instead, computation of the diffusion update equation results in a filtered representation where small objects of high contrast may be smoothed.

The construction of the scale generating linear filter will be developed in the following sections, and in this discussion, design criteria of the linear filter will be described by relationships between specific kernel values. The kernel of the linear filter will be referred to as $H[\cdot]$, and individual components of the kernel will be described by indices within the kernel vector. For example, the influence of a gradient in its own decision to diffuse is defined by the center value of the linear kernel, denoted as $H[0]$.

Analyzing the suitability of incorporating a specific linear filter into
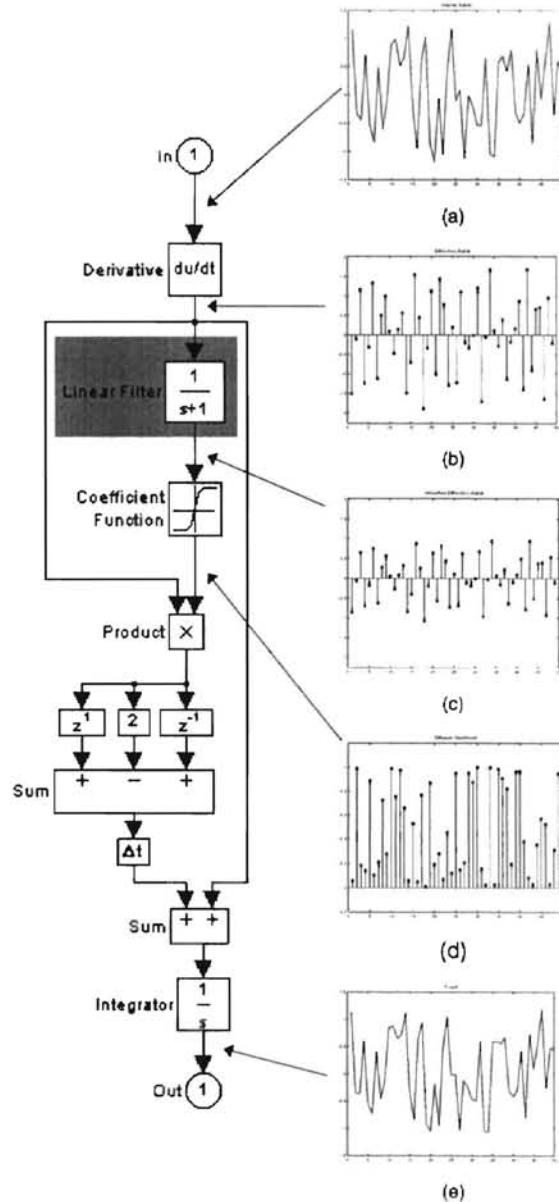


**Figure 3.4** A system representation for one iteration of the scale aware anisotropic diffusion process: (a) the input signal, (b) its difference representation, (c) prefiltered with the linear, scale generating filter, (d) diffusion coefficients, and (e) the filtered result. The only difference between a scale aware diffusion expression and a traditional representation is the addition of the linear filter, denoted with the gray box.

35

the diffusion coefficient initially imposes no restrictions on the construction of a linear filter. Theoretically, any kernel could be convolved with the signal gradient before the coefficient calculation, and the discussion in the following chapters only assumes a candidate linear filter to be symmetric, as the spatial smoothing properties of a scale aware diffusion expression should smooth small scale object independent of their spatial relationships. However, in reviewing linear kernels utilized in the literature, FIR lowpass filters are often chosen to increase the scope of the coefficient calculation, and with the purpose of providing an intuitive representation of the system, three example filter kernels are presented in Figure 3.5.
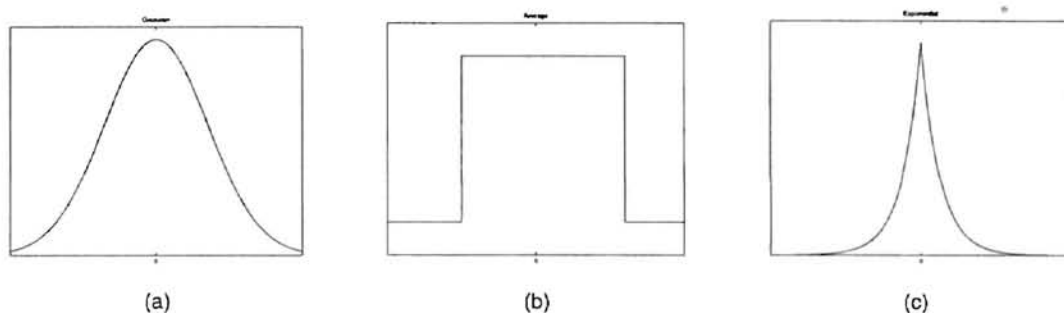


**Figure 3.5** Three example kernels which could be incorporated into a spatially aware anisotropic diffusion coefficient: (a) Gaussian, (b) Average, and (c) Exponential filters.

## Generalized Anisotropic Diffusion

Following the function definitions of the previous section, this section seeks to develop a general representation of the anisotropic diffusion equation. Several diffusion coefficient formulations have previously been presented, and the purpose of this section is

to establish conditions under which all diffusion expressions have similar performance properties. Specifically, this section will explore cases guaranteeing the preservation of an edge, regardless of the specific realization of the anisotropic diffusion coefficient. This generalized model will then be used in the next section to show that the spatially aware anisotropic diffusion equation is unable to simultaneously smooth regions of low contrast and small spatial size.



(a)



(b)
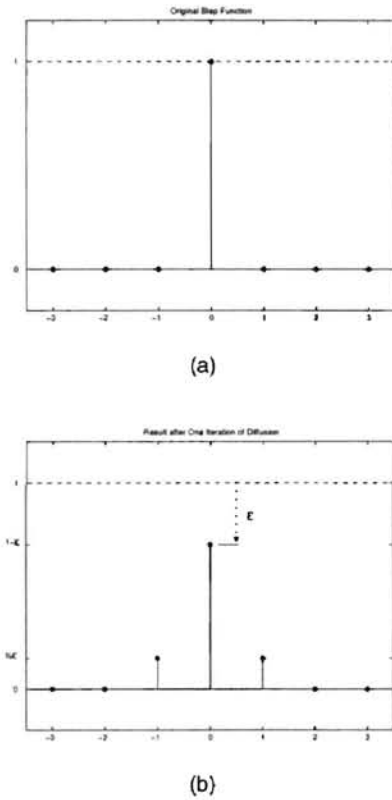
**Figure 3.6** Gradient representations of a step function: (a) the original step function and (b) the result after smoothing with one iteration of anisotropic diffusion.

The first step in dividing the analysis of the anisotropic diffusion smoothing filter from the implementation of its diffusion coefficient is to consider the characteristic smoothing of the step function. The difference in the original step function, before and after one iteration of smoothing, is shown in Figure 3.6, where the change of the signal gradient between iterations is given as

$$\varepsilon = \nabla \mathbf{I}_{t+\Delta t} - \nabla \mathbf{I}_t = \Delta t \left[ -2\mathbf{c}(\nabla \mathbf{I}_t) \cdot \nabla \mathbf{I}_t \right] \quad (3.8)$$

using (3.7).

While the manner in which the step function will smooth differs dependent on the realization of the anisotropic diffusion coefficient, all coefficient expressions are defined to have the common form of decreasing with increasing gradient, in order to realize their goal of preserving regions of possible transition. Independent of the anisotropic diffusion

37

expressions, as the step function increases in gradient magnitude, the diffusion coefficient is defined to decrease towards zero and will inhibit smoothing at possible edge locations; therefore, step functions with large gradient will sustain little change between iterations and will be preserved by the diffusion operation.

The minimal height of a step edge which will remain unsmoothed by the anisotropic diffusion equation will be referred to frequently in this thesis, and it is defined as having a magnitude greater than $\beta(K_{DC},t)$, a function which denotes the *smoothing threshold* of the diffusion system on the step edge. Gradients below $\beta(K_{DC},t)$ will have significant values of $\varepsilon$ and will be smoothed. Gradients above $\beta(K_{DC},t)$ will undergo little change between iterations ($\varepsilon \cong 0$) and will be preserved. The two parameters of the smoothing threshold correspond to the variables of the anisotropic diffusion operation. The first parameter, $K_{DC}$, is analogous to the gradient threshold, $k$, found in all diffusion coefficients, with the subscript $DC$ denoting that the gradient thresholds of different coefficients are not directly comparable. The second parameter, $t$, is the solution time or scale of the anisotropic diffusion process, indicating that small changes of the signal gradient will be amplified by increasing solution time.

Smoothing thresholds are initially defined in terms of step functions, but the general smoothing performance of the anisotropic diffusion equation is not completely described by its response to the step function (unlike its Gaussian kernel parent). While the smoothing threshold is not valid for modeling the diffusion of an arbitrary initial sequence, it is a representative model for smoothing a restricted class of input signals. By considering an arbitrary sequence of step functions, it can be proven that a gradient

38

greater than $\beta(K_{DC},t)$ is guaranteed to be preserved if it is within a signal region which is locally monotonic. This property is formally derived in Appendix B and a visual example is presented in Figure 3.7.

The development of a smoothing threshold abstracts the anisotropic diffusion equation from its coefficient realization, guaranteeing the preservation and identification of certain features in all diffusion implementations. The only requirement on the diffusion coefficient is that it decrease with increasing gradient, tending towards zero as gradients become large. In the next subsection these concepts are applied to the study of the scale aware diffusion process and its capability of removing small scale features.

(a)

(b)
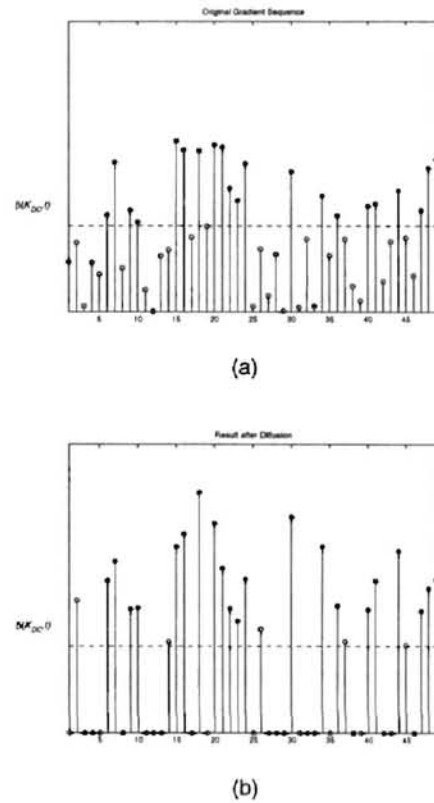
**Figure 3.7** A visual example of the smoothing performance of traditional anisotropic diffusion within a monotonic region. Gradient representations of (a) the original signal and (b) its filtered result, notice that all gradients with magnitude greater than $\beta(K_{DC},t)$ are not diminished by the filtering operation.

Incorporating a linear filter into the diffusion coefficient calculation allows the removal of small scale features and the possibility of constructing an image pyramid utilizing the anisotropic diffusion equation. In this section the result of using an arbitrary linear filter in the diffusion process is explored. The section begins with analysis of the diffusion equation applied to the finite width step function, deriving conditions on the linear filter which guarantee the removal of these small scale regions. A sequence of step functions will then be used to develop constraints which allow the diffusion equation to realize its second design objective, smoothing regions of small contrast. The section ends by showing the duality of these requirements, concluding that a diffusion coefficient incorporating any linear filter will be incapable of simultaneously removing small regions and preserving edges. Therefore, traditional scale aware diffusion processes are unsuitable for image pyramid construction.

Use of a linear filter within the diffusion coefficient necessitates a clarification of the smoothing threshold presented in the previous section. Traditional diffusion coefficients are calculated from the underlying gradients of the original signal, and defining a step edge with height greater than the smoothing threshold, $\beta(K_{DC},t)$, denotes that the resulting diffusion coefficient will be near zero and that the step function will be preserved. Incorporating a linear filter into the diffusion expression modifies the gradients used in the coefficient calculation, and defining a step function of height $\alpha$ to be greater than $\beta(K_{DC},t)$ does not account for the effect of the linear filter on the coefficient

calculation and does not guarantee the preservation of the feature. When using a linear filter within the diffusion coefficient, a step function will only be preserved if the gradient used in the calculation of the diffusion coefficient, $|\alpha\mathbf{H}[0]|$, is greater than the smoothing threshold, $\beta(K_{DC},t)$, where $\mathbf{H}[0]$ is the kernel value of the linear filter at $x = 0$, and in the rest of this chapter, defining a step function to have magnitude greater than the smoothing threshold indicates that $|\alpha\mathbf{H}[0]|\geq\beta(K_{DC},t)$.

Spatial Smoothing

Differences between the traditional anisotropic diffusion equation and the spatially aware extensions are evident in the smoothing of a finite width step function with height $\alpha$. If the height of the function is larger than $\beta(K_{DC},t)$, the traditional diffusion equation will be unable to smooth either gradient and will maintain the feature. The goal of a scale aware anisotropic diffusion process is to smooth all regions with small spatial size, independent of gradient magnitude.

In describing a class of linear filters which, when incorporated into the diffusion coefficient, would allow the anisotropic diffusion equation to remove the finite width step function of width $W$, it is required that the filtered gradients used in the coefficient calculation be smaller than the smoothing threshold, $\beta(K_{DC},t)$. Mathematically, the scale aware diffusion equation must satisfy

$$\left|\alpha(\mathbf{H}[0]-\mathbf{H}[W])\right| < \beta(K_{DC},t) \tag{3.9}$$

when

$$\left|\alpha\mathbf{H}[0]\right| \geq \beta(K_{DC},t), \tag{3.10}$$

41

where $\alpha$ is the height of the finite width step function, $W$ is its width, $\alpha\mathbf{H}[0]$ is the contribution of the positive gradient in the decision to smooth at location 0, $-\alpha\mathbf{H}[W]$ is the contribution of the negative gradient in the decision to smooth at location 0, and $\alpha(\mathbf{H}[0]-\mathbf{H}[W])$ is the gradient seen by the coefficient calculation at location 0. These two conditions define a diffusive process which will remove a finite width step function determined solely by its scale.

Candidate linear filters are further constrained by examining all step functions with width less than the spatial smoothing goal of the anisotropic diffusion expression, $W$. These functions should also be removed by a scale aware diffusion process, and this is guaranteed when

$$\left|\alpha(\mathbf{H}[0]-\mathbf{H}[n])\right| < \beta(K_{DC},t) \qquad 0 < n \le W, \tag{3.11}$$

where $n$ is the width of the finite width step function and $\alpha(\mathbf{H}[0]-\mathbf{H}[n])$ is the gradient seen by the coefficient calculation at location 0.

Initial requirements on the linear filter for smoothing regions of small spatial extent appear to be $\mathbf{H}[0] > (\mathbf{H}[0]-\mathbf{H}[n])$; however, the presence of the gradient magnitude, $\alpha$, in (3.9) and (3.11) couples the spatial performance of the diffusion equation to the height of the feature. Ideally, the spatially aware anisotropic diffusion equation would be capable of removing small regions, independent of their gradient magnitudes. As the feature intensity becomes larger than $\beta(K_{DC},t)$, $\alpha$ becomes large and the difference between $\mathbf{H}[0]$ and $\mathbf{H}[n]$ is amplified. If $\mathbf{H}[0] \neq \mathbf{H}[n]$, the limit

$$\lim_{\alpha \to \infty} \left|\alpha(\mathbf{H}[0]-\mathbf{H}[n])\right| \ge \beta(K_{DC},t), \tag{3.12}$$

42

guarantees that the spatially aware diffusion equation will be incapable of smoothing a finite width step function with significant magnitude. Setting $\mathbf{H}[0] \cong \mathbf{H}[n]$, removes the gradient magnitude of the finite width step function from the diffusion decision, and since

$$|\alpha \cdot 0| < \beta(K_{DC}, t) \quad \forall \alpha, \tag{3.13}$$

a diffusion expression incorporating the class of linear filters of the form

$$\mathbf{H}[0] \cong \mathbf{H}[1] \cong \cdots \cong \mathbf{H}[W-1] \cong \mathbf{H}[W] \tag{3.14}$$

will be capable of removing all small regions of high contrast.

While analysis of the finite width step function develops conditions on linear filters which allow an anisotropic diffusion process to remove regions of small spatial size, the next subsection will show that these conditions greatly handicap diffusion from smoothing features of low contrast. Conditions on the linear filter which allow the smoothing of low contrast regions will be developed in the next section, accomplished through analysis of two step functions. It will be shown that the resulting design criteria does not allow the construction of an anisotropic diffusion operation which, through incorporating a linear filter into its coefficient calculation, is capable of removing small objects while smoothing regions of small contrast.

## Gradient Smoothing

Traditional anisotropic diffusion expressions smooth regions of small contrast and preserve candidate edge locations. An ideal scale aware anisotropic diffusion operator should attempt to mimic the traditional diffusion processes in the absence of small features. Consider the smoothing performance of anisotropic diffusion on the sequence

43

shown in Figure 3.8, consisting of two step functions of height $\gamma$ and $\alpha$, respectively. Letting the gradient magnitude of the first step function be small ($|\gamma\mathbf{H}[0]| \cong 0$) and the second gradient magnitude be large ($|\alpha\mathbf{H}[0]| \geq \beta(K_{DC}, t)$), traditional anisotropic diffusion would preserve (and possibly enhance) the second step edge while smoothing the smaller one.

Presence of a linear filter within a diffusion coefficient allows an anisotropic diffusion process to remove small regions. In the sequence presented above, no small features are present and the spatially aware anisotropic diffusion equation should produce smoothing equivalent to the traditional anisotropic diffusion process, smoothing the smaller step function while preserving the larger. Mathematically, the larger step function will be preserved by the spatially aware anisotropic diffusion expression if

$$|\alpha\mathbf{H}[0] + \gamma\mathbf{H}[-n]| \geq \beta(K_{DC}, t), \quad (3.15)$$



(a)

(b)

**Figure 3.8** A sequence of two step functions with arbitrary height. The equivalent gradient representation is shown in (b).

where $\mathbf{H}[\cdot]$ is the linear kernel incorporated into the diffusion coefficient and $\alpha\mathbf{H}[0] + \gamma\mathbf{H}[-n]$ is the gradient seen by the coefficient calculation at location $n$. Since $|\alpha\mathbf{H}[0]|$ was initially defined to be greater than the smoothing threshold and $\gamma$ was

defined to be positive, preservation of the larger gradient is assured if the filter coefficients of $\mathbf{H}[\cdot]$ have the same sign.

Conditions that guarantee the removal of the smaller step functions introduce further constraints on linear filter construction. Traditional anisotropic diffusion would smooth the smaller step function, and in a spatially aware anisotropic diffusion process, the smoothing of this step function will occur when

$$\left|\gamma\mathbf{H}[0]+\alpha\mathbf{H}[n]\right| < \beta\left(K_{DC},t\right), \tag{3.16}$$

where $\gamma\mathbf{H}[0]+\alpha\mathbf{H}[n]$ is the gradient seen by the coefficient calculation at location 0. Though not sufficient for assuring the smoothing of the smaller step function, a necessary condition for its removal is

$$\left|\alpha\mathbf{H}[n]\right| < \beta\left(K_{DC},t\right), \tag{3.17}$$

where $\alpha\mathbf{H}[n]$ is the influence of the larger step function on the smoothing of the smaller.

Therefore, removal of the smaller step function is facilitated when $\left|\alpha\mathbf{H}[n]\right|$ is less than the smoothing threshold, while it has previously been defined that $\left|\alpha\mathbf{H}[0]\right|$ is greater than the smoothing threshold. Candidate diffusion coefficients, capable of allowing the smoothing of small gradients, are initially constrained by these two conditions to incorporate linear filters of a form $\mathbf{H}[0] > \mathbf{H}[n]$. Linear filters suitable for coefficient construction are further constrained by considering the effect of gradient magnitude on the smoothing performance. Traditional anisotropic diffusion would smooth the smaller step function regardless of the amplitude of the larger step function, and attempting to construct a scale aware diffusion coefficient capable of mimicking this behavior further

45

restricts the candidate filters. Allowing $\alpha$ to increase in magnitude, while still requiring the diffusive process to smooth the smaller step function, requires a diffusion coefficient incorporating a linear filter of the form $\mathbf{H}[0] \gg \mathbf{H}[n]$.

This conflicting condition realizes the goal of this chapter, showing that an anisotropic diffusion coefficient utilizing a linear filter to increase its scale is unable to simultaneously remove regions of small spatial size and low contrast. Study of the finite width step function has shown that the successful removal of small objects with the anisotropic diffusion equation is guaranteed when the difference between the kernel values of the linear filter approach zero ($\mathbf{H}[0] \cong \mathbf{H}[n]$). Conversely, study of the second sequence of step functions has displayed that the characteristic anisotropic smoothing of a signal, preserving features of high gradient while smoothing smaller gradient features, is imitated by a scale aware diffusion process as the kernel values within the linear filter decrease sharply ($\mathbf{H}[0] \gg \mathbf{H}[n]$). Observing the contradiction of these two conditions, this chapter may conclude that it is impossible to design a linear filter which, when incorporated into an anisotropic diffusion process, will allow the systematic removal of small scale features and the simultaneous removal of regions of low contrast.

Presence of the gradient magnitudes within the coefficient calculation results in conditions that make it impossible to design a linear filter capable of identifying small scale objects while still allowing the diffusion equation to smooth regions of small gradient. Construction of an anisotropic diffusion pyramid requires a smoothing operator capable of maintaining edge locations and producing filtered representations suitable for sampling, and in developing a scale aware diffusion expression possessing these

characteristics, studies of the linear filters suggest that gradient magnitudes should be separated from the spatial smoothing goals of the diffusion equation. This requirement motivates the use of a nonlinear filter, and in the next section, nonlinear morphological filters will be introduced into the anisotropic diffusion coefficient calculation. It will be shown that the resulting diffusive process is capable of smoothing small objects independent of intensity and without projecting spatial influence. These filter characteristics describe a filtering process suitable for image pyramid construction.

# CHAPTER IV

# MORPHOLOGICAL ANISOTROPIC DIFFUSION

## Overview

Anisotropic diffusion was originally designed to generate scale spaces with minimal feature movement. Construction of scale spaces with this property would alleviate problems in applying coarse to fine search methods to image pyramids. Image pyramids, however, require a scale generating process which guarantees the removal of information and satisfies necessary sampling conditions, and while traditional anisotropic diffusion expressions are incapable of generating these filtered results, scale aware extensions of the anisotropic diffusion equation attempt to smooth regions of low contrast and small spatial size. A smoothing operator possessing these characteristics would generate a filtered result suitable for sampling and an image pyramid suitable for coarse to fine search methods.

Initial scale aware realizations of the anisotropic diffusion equation incorporate linear filters into the diffusion coefficient and effectively increase the scope of the coefficient calculation, allowing the smoothing of small regions. The previous chapter has shown that these original scale aware diffusion expressions are severely handicapped, as they are unable to smooth regions of small gradient while removing small scale features, and analysis of their smoothing performance implies that development of a scale

aware diffusion expression requires regions to be identified without removing their high frequency information.

This criterion suggests the use of nonlinear filters in increasing the scope of the coefficient calculation and allowing the diffusion expression to smooth objects of small spatial scale. The purpose of this chapter is to explore anisotropic diffusion processes which incorporate nonlinear morphological filters into their diffusion coefficients, and the goal of this chapter is to show that these new diffusion expressions are capable of smoothing features of small spatial size while simultaneously smoothing regions of low contrast. This will be accomplished through both theoretical and experimental analysis and will allow the new morphological anisotropic diffusion processes to be utilized in the following chapters, as the construction of an anisotropic diffusion pyramid may then be formally defined.

This chapter is organized as follows: Section II introduces morphological filters and develops their equivalent gradient representations. Section III derives conditions on the morphological operators necessary for smoothing regions of small spatial size and smoothing areas of low contrast, and shows that these two conditions may be simultaneously satisfied by a class of morphological filters. Section IV presents experimental results and compares the smoothing characteristics of the new diffusion coefficient to traditional and scale aware anisotropic diffusion realizations.

Morphological Filtering

Morphological operators are able to produce image representations of increasing scale without eradicating edges. Approaching image processing from the vantage point of human perception, morphological operators simplify image data, preserving essential shape characteristics and eliminating irrelevancies [33, 17]. The purpose of this section is to provide an introduction to these nonlinear filters, with emphasis on developing an equivalent gradient domain representation of their smoothing characteristics. This discussion will utilize the fundamental edge models of the previous section and will conclude with the incorporation of a candidate morphological filter into a diffusion coefficient.

Mathematical morphology was originally developed in the context of algebraic topology and is derived from the investigation of shape. Its use in digital signal processing is defined with two fundamental operators - *erosion* and *dilation*. Implementation of the erosion and dilation filters is similar to a median filter and is accomplished with nonlinear minimum and maximum operations. An erosion removes regions of high intensity and is expressed as

$$\mathbf{I} \ominus \mathbf{M} = \min_{y \in \mathbf{M}} \{ \mathbf{I}(x - y) \}, \tag{4.1}$$

where $\mathbf{M}$ is the structuring element and $\ominus$ is the erosion operator. The mathematical dual of erosion is dilation and removes regions of low contrast by computing the maximum value within a region. A dilation is expressed as

$$\mathbf{I} \oplus \mathbf{M} = \max_{y \in \mathbf{M}} \{ \mathbf{I}(x - y) \}, \tag{4.2}$$

50

where $\oplus$ is the dilation operator. In both fundamental filters, the realization of the structuring element defines the shape and scale of the morphological filter and conceptually denotes the signal region from which the minimum (or maximum) value is drawn.

Developing an equivalent gradient representation of the morphological operators is a major goal of this section and begins with analysis of the step function. These functions have been shown to be eigenfunctions of the morphological system[†] [24], and the effect of the morphological filters on the step function is shown in Figure 4.1. Notice that the resulting functions are not smoothed representations of the signal and, instead, are simply shifted by half the width of the structuring element. Applying an erosion to the edge function translates the signal to the right. Alternatively, filtering with a dilation shifts it to the left. The sequential application of these filters results in an infinite number of paths for the step function gradient to travel, but never modifies the gradient magnitude. Figure 4.1b displays the right-hand derivatives of the function and its filtered results, again showing the movement of the edge.

---

[†] A morphological eigenfunction remains unchanged in signal amplitude and structure after being filtered with a morphological operation. The only effect of the morphological filter on the eigenfunction is a possible translation of the original signal.
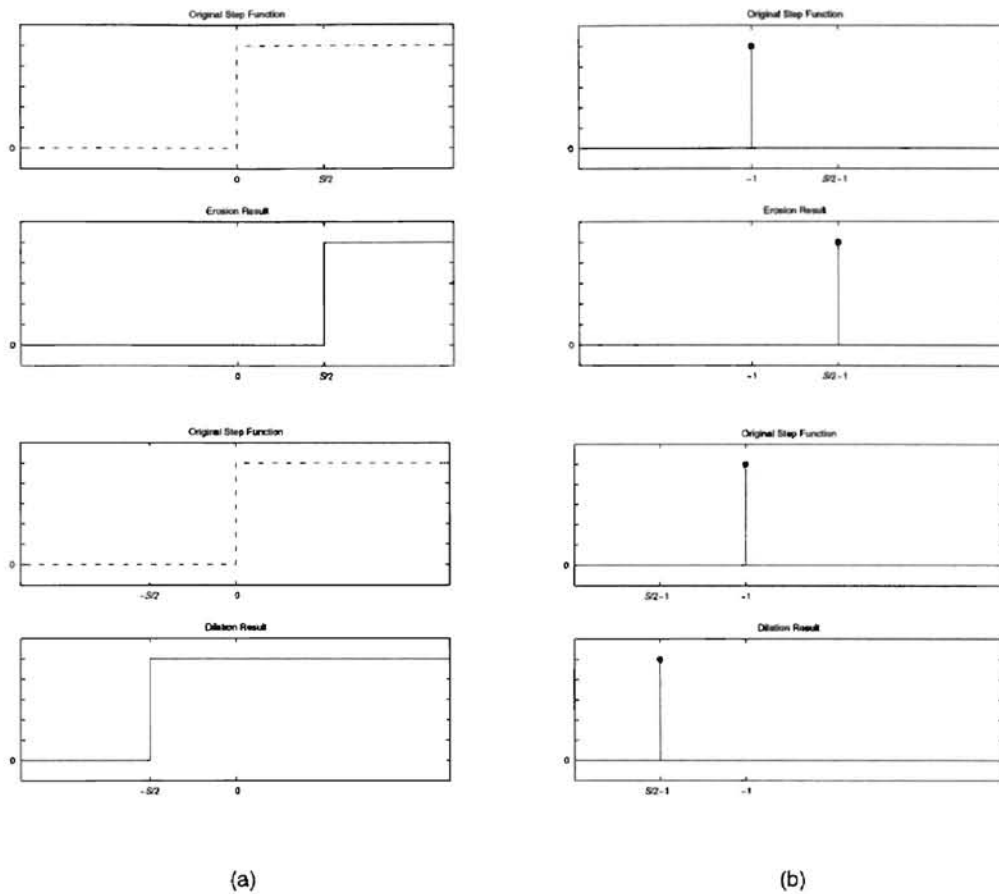
**Figure 4.1** Effect of the fundamental morphological filters on a step function. The equivalent gradient representations are shown in (b).

A complete gradient understanding of the morphological operators continues with investigation of the negative of the original step function. The signal and its filtered results are displayed in Figure 4.2. Again, the morphological filters produce translation of the step edge and do not effect the signal amplitude. An important observation is that while the erosion translates the positive edge of Figure 4.1 to the right, it translates the negative edge in Figure 4.2 to the left. The dilation produces a similar result, transporting positive and negative edges in opposite directions. This property of sign dependent translation defines the performance of the morphological filters.
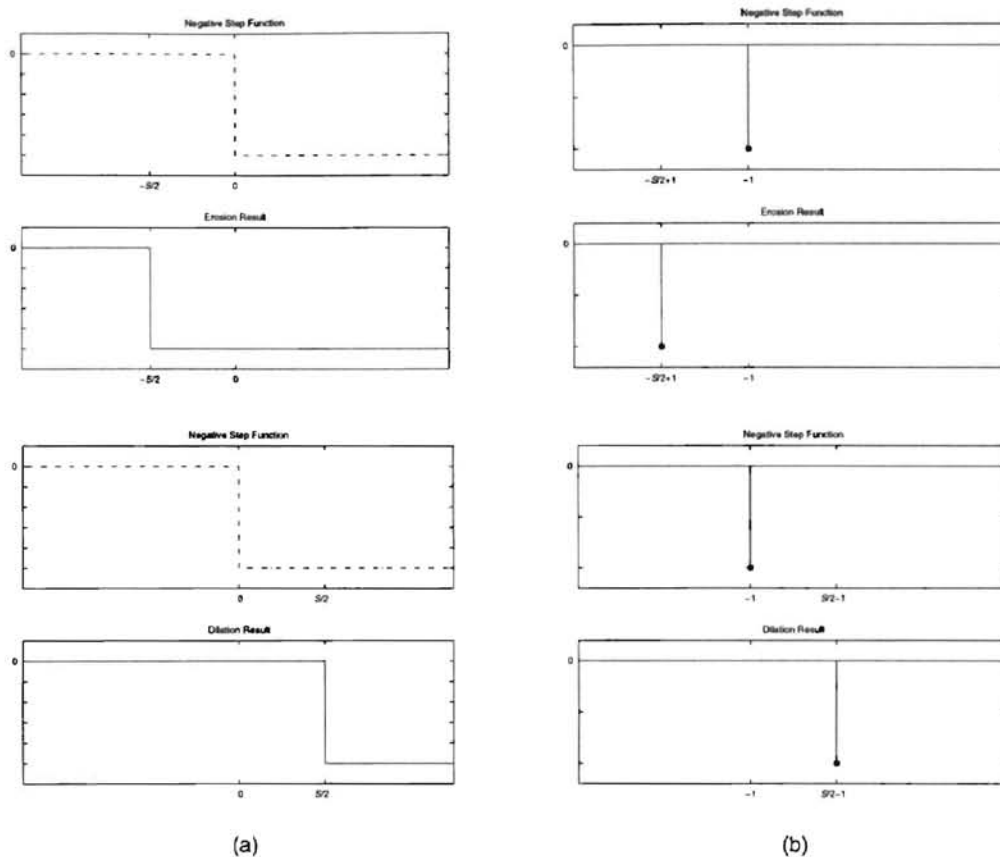
**Figure 4.2** Effect of the fundamental morphological filters on a negative step function. The equivalent gradient representations are shown in (b).

Morphological filtering occurs as negative and positive gradients interact and remove each other. As displayed with the finite width edge model, eroding the function will eliminate it as the structuring element of the morphological filter becomes large. Results of filtering the finite width edge function using an erosion operation with several different structuring element sizes are shown in Figure 4.3, and equivalent gradient representations are shown in Figure 4.3b.
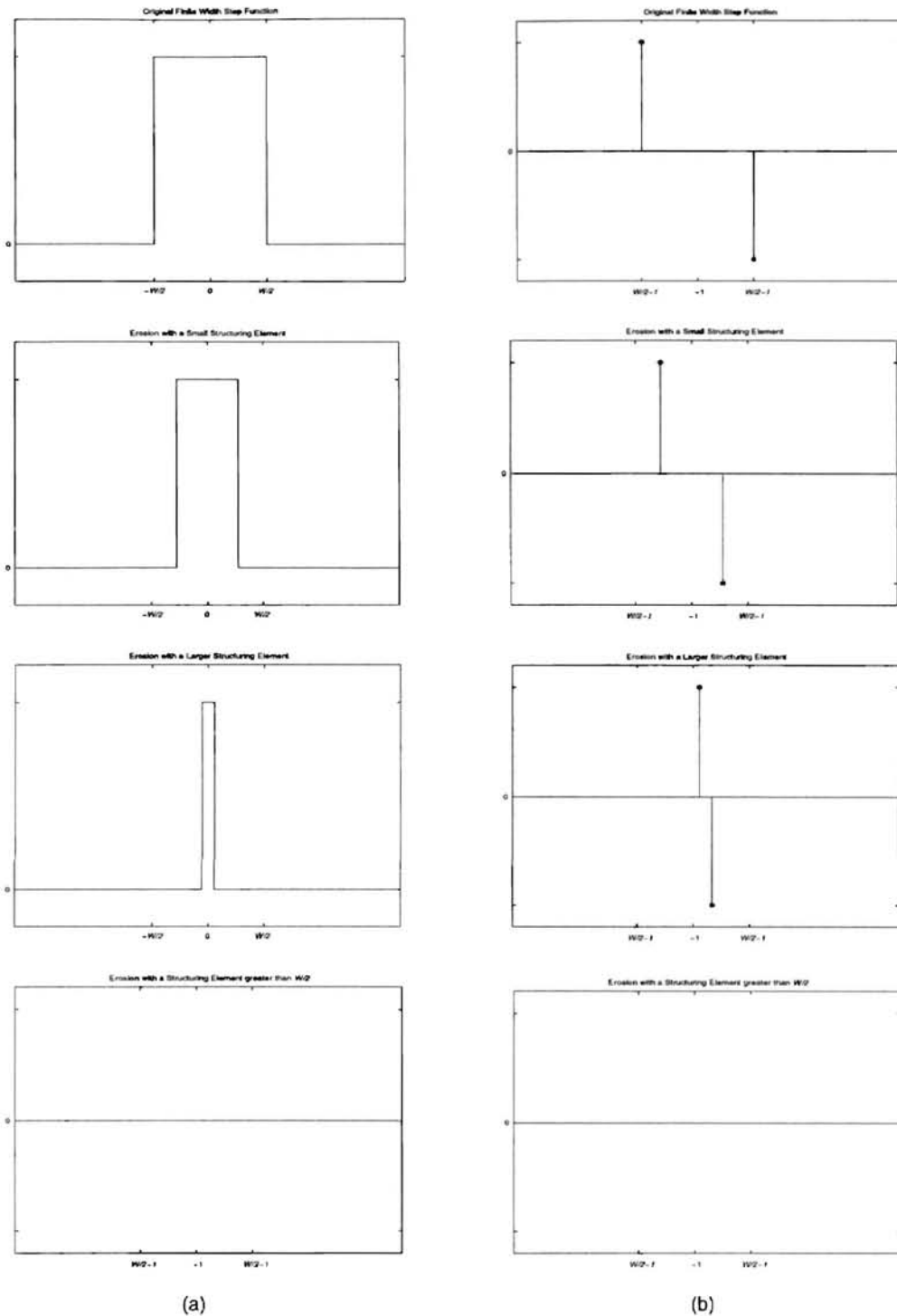
**Figure 4.3** Eroding a finite width step function with structuring elements of different size. From top to bottom: the original finite width step function; the finite width step function eroded with a small structuring element; the finite width step function eroded by a larger structuring element; the finite width step function eroded with a structuring element larger than W/2. The equivalent gradient representations are shown in (b).

54

Analysis of the morphological operators becomes more difficult to visualize with an arbitrary signal. Relying on the eigenfunction properties of the step function, the sign dependent gradient representations, presented above, define the performance of the fundamental morphological filters with respect to edge preservation and movement. An erosion may be expressed as

$$\nabla \mathbf{I}_{t+\Delta t}(x) = \nabla^+ \mathbf{I}_t \left( x - \frac{\Delta t}{2} \right) + \nabla^- \mathbf{I}_t \left( x + \frac{\Delta t}{2} \right), \tag{4.2}$$

where $t+\Delta t$ is the width of the structuring element, $\nabla \mathbf{I}_t$ is the original image, $\nabla^+$ is the maximum value of either the gradient or zero, $\nabla^-$ is the minimum value of either the gradient or zero, and $\Delta t$ is the time step. A dilation is realized by reversing the direction of gradient propagation and expressed as

$$\nabla \mathbf{I}_{t+\Delta t}(x) = \nabla^+ \mathbf{I}_t \left( x + \frac{\Delta t}{2} \right) + \nabla^- \mathbf{I}_t \left( x - \frac{\Delta t}{2} \right). \tag{4.3}$$

For discrete implementations of these fundamental morphological filters, $\Delta t$ should be one.

Application of a single erosion or dilation operation removes regions dependent on intensity. It also results in edge movement. In the anisotropic diffusion equation, it is of interest to remove regions independent of intensity and without inducing edge movement, and the sequential application of the fundamental filters can produce morphological systems realizing these goals. An *open* operation removes regions of low intensity, without introducing feature drift, and is implemented by first eroding a signal and then dilating the result. The *close* operation removes regions of high intensity, without inducing feature drift, and is implemented by dilating and eroding. Combinations

of these higher level processes remove regions independent of intensity and without causing edge movement.

Introducing morphological operators into the diffusion coefficient is straightforward and concludes this section. Using the coefficient presented in (2.5) as an example, a possible realization of a morphological scale aware diffusion coefficient is

$$c(\nabla \mathbf{I}) = e^{-\left(\frac{\|\nabla(\mathbf{I} \oplus \mathbf{M})\|}{k}\right)^2}, \tag{4.4}$$

where $\mathbf{I} \oplus \mathbf{M}$ is the erosion of the signal $\mathbf{I}$ with structuring element $\mathbf{M}$. The choice of morphological filter type and structuring element size is crucial to the construction of a scale aware diffusion coefficient and it will be the focus of the next section, where conditions will be developed which define morphological operators capable of identifying objects of small spatial size without interfering in the smoothing of small gradients.

Theoretical Analysis

The purpose of this section is to show that incorporating morphological filters into the anisotropic diffusion equation *does* create a process which can remove features based solely on gradient or spatial scale and is, therefore, *applicable* to image pyramid construction. This section parallels the analysis presented of the linear filters in the previous chapter and provides investigation of the smoothing performance of the morphological filters on the step and finite width edge models. All analysis is accomplished using the previously defined gradient representations of the morphological filters and the anisotropic diffusion equation.

## Spatial Smoothing

The spatial smoothing of a morphological anisotropic diffusion system is displayed with the finite width step function. Considering a step function whose amplitude is greater than the smoothing threshold, $\beta(K_{DC},t)$, traditional anisotropic diffusion expressions will preserve the gradients and maintain the feature. The goal of a scale aware diffusion process is to remove the region, independent of gradient magnitude.

In describing a morphological filter which, when incorporated into the diffusion coefficient, will smooth the finite width step function, it is necessary to define two properties of the morphological filtering operation: the direction of gradient propagation and the distance of gradient translation. The first characteristic, the direction of gradient propagation, is defined by the morphological filter type and denotes whether positive gradients are shifted to the right or left. (Negative gradients will be shifted in the opposite direction.) The second characteristic, the distance of gradient translation, describes the spatial distance over which the gradients are moved and is defined by the solution time of the morphological system or, equivalently, the structuring element size.

Using a morphological filtering operation to smooth a finite width step function provides an initial description of morphological sequences suitable for incorporating into the diffusion equation, and filtering the finite width step function requires that the gradients of the edge model interact and remove each other. Remembering that the finite width step function is constructed with a positive gradient located at the origin and a negative gradient located at location $W$, where $W$ is the width of the edge model, feature

removal necessitates that the gradients move towards each other and requires that the positive gradient must move to its right while the negative gradient travels to its left.

Construction of the finite width step function defines the direction of gradient propagation necessary for filtering, and interaction between the two gradients of the edge model defines the distance of gradient translation. Removal of the finite width edge function will occur when the positive and negative gradients interact, and since the gradients move towards each other with equal speed, their interaction will occur at the center of the edge model. Specification of a morphological operation which allows the diffusion equation to smooth small objects requires a morphological filtering sequence which shifts positive gradients to the right a distance of $W/2$. This requirement is shown graphically in Figure 4.4, and an example morphological sequence satisfying these requirements is shown in (4.2), described by an erosion operation. Solved for solution times greater than $W$, the erosion operation will translate positive gradients to the right a distance of $W/2$ and is a candidate filter for inclusion within a scale aware diffusion coefficient.
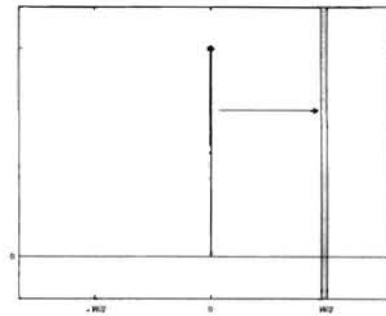


**Figure 4.4** Necessary gradient movement for removing a finite width step function of width $W$ with a morphological filter.

Scale aware diffusion expressions should remove regions of width $W$, and the erosion operation presented above satisfies this requirement. Scale aware diffusion expressions should also remove regions smaller than the spatial smoothing goals of the

58

anisotropic diffusion equation, and analysis of smaller finite width edge models develops conditions necessary for their removal. Smoothing of a finite width step function with width $n$ ($0<n\leq W$) still requires positive gradients be transported to the right, but only necessitates that they be translated over a distance of $n/2$. The requirements for removing all finite width step functions of width $W$ or less are shown graphically in Figure 4.5, and an example morphological sequence capable of satisfying these conditions is the erosion operation solved solution times greater than $W$ - the identical filtering operation derived for smoothing the larger finite width step functions.
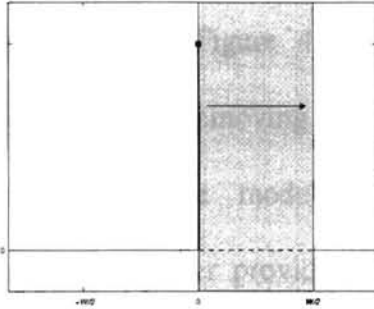


**Figure 4.5** Necessary gradient movement for removing all finite width step functions of width less than or equal to $W$ with a morphological filter.

While study of the smaller finite width edge model did not refine the requirements of candidate morphological sequences that are suitable for application to the anisotropic diffusion coefficient, analysis of the negative finite width step function does further constrain the construction of the morphological filter and motivates the need for a more complex filter sequence. The negative finite width step function is given as

$$\mathrm{u}_{fw}(x) = \alpha \mathrm{u}(n) - \alpha \mathrm{u}(0), \quad 0 < n \leq W \tag{4.5}$$

and should also be removed by the scale aware diffusion equation. Smoothing of this function introduces different requirements on the direction of gradient propagation within the morphological filter sequence, as the positive gradient of this edge model is located to the right of the negative gradient. (The positive gradient of the original finite width step

59

function was located to the left of the negative gradient.) Filtering necessitates that the gradients move towards one another, and for the negative finite width step function requires that the positive gradient travel to its left and the negative gradient travel to its right. Feature removal will occur when the gradients interact at the center of the edge model, $n/2$, and Figure 4.6 summarizes the requirements for removing all of the negative finite width edge models. A candidate morphological filter providing this smoothing is the dilation operation presented in (4.3) and solved for solution times greater than $W$.



**Figure 4.6** Necessary gradient movement for removing all negative finite width step function of width less than or equal to $W$ with a morphological filter.

Consolidating the requirements for removing both finite width step functions concludes this subsection and defines the class of morphological filters suitable for providing a diffusion coefficient with the capability of identifying and smoothing regions of small spatial size. It has been shown that removing a positive finite width step function requires a
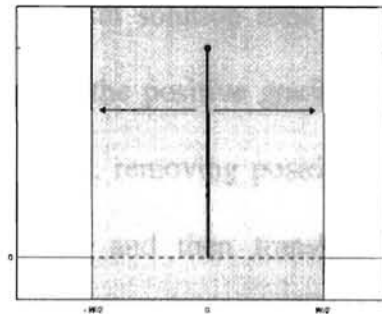


**Figure 4.7** Necessary gradient movement for removing all negative and positive finite width step functions of width less than or equal to $W$ with a morphological filter.

morphological filter sequence which translates positive gradients to the right and that smoothing a negative finite width step function requires a filtering operation capable of translating positive gradients to the left. In both smoothing examples, the original

60

gradients must be transported over a distance equal to half the spatial smoothing goals of the anisotropic diffusion equation, $W/2$. These requirements are summarized graphically in Figure 4.7.

While the single application of an erosion or dilation operation will not generate the necessary smoothing performance for filtering both edge models, the sequential concatenation of these fundamental morphological operators will produce a filter capable of satisfying these requirements, transporting gradients throughout the desired regions. Many morphological operations could be constructed, and an example morphological filter sequence, which produces the necessary gradient movement, is shown graphically in Figure 4.8. The sequence is realized by dilating the signal with solution time $W+1$ and then eroding the result with solution time $2(W+1)$. This initially moves the positive gradient over the region to the left, removing positive finite width edge models, and then translates the gradient back to the origin and through the region to the right, removing negative finite width edge models.



**Figure 4.8** Necessary gradient movement for removing all negative and positive finite width step functions of width less than or equal to *W*. The morphological sequence consists of: (a) dilation, (b) erosion, and (c) erosion.

Analysis of the finite width step functions defines a class of morphological operators which remove features of small scale and whose incorporation into a diffusion coefficient would allow the anisotropic diffusion expression to smooth these small scale regions. The purpose of the next subsection is to define a class of morphological filters

which allows the anisotropic diffusion equation to smooth regions of low contrast. After deriving these smoothing conditions, the section will conclude that incorporating certain morphological filter sequences into a diffusive process develops a smoothing operation capable of simultaneously removing objects of small spatial size while smoothing gradients of small scale.

Gradient Smoothing

Filtering of small scale objects necessitates that positive and negative gradients interact, motivating morphological operators to be described through the regions over which a gradient must travel. Smoothing of small contrast regions with the anisotropic diffusion equation requires that significant gradient magnitudes not interfere in diffusion, and in describing morphological filters which do not inhibit the smoothing of low contrast areas, filter types must only defined by the net distance of gradient movement and are unconcerned with the specific path a gradient undertakes. Consider the sequence of two step functions shown in Figure 4.9 and previously presented in the last chapter. Application of morphological operators to this sequence will



(a)



(b)
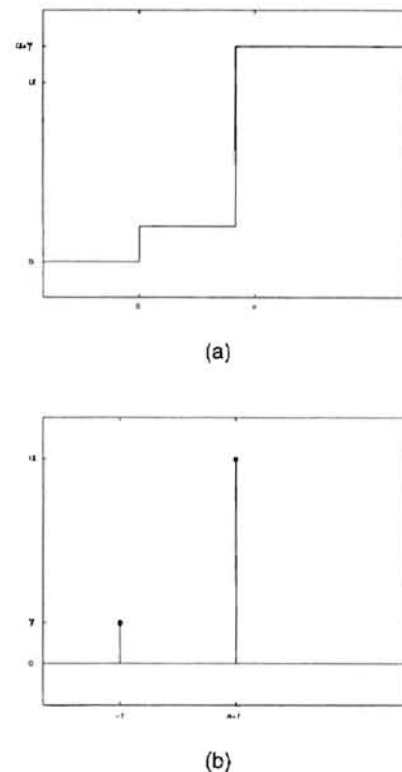
**Figure 4.9** A sequence of two step functions with arbitrary height. The equivalent gradient representation is shown in (b).

62

never change the structure of the signal, as both gradients will travel in the same direction and at the same speed. Morphological operators applied to monotonic regions can only introduce delay.

Defining the height of the first step function in Figure 4.9 to have small magnitude ($\gamma \cong 0$) and the height of the second step function to be significant ($\alpha \geq \beta(K_{DC}, t)$), traditional anisotropic diffusion expressions will maintain the larger edge function while smoothing the smaller step. An ideal scale aware diffusion expression will attempt to reproduce the smoothing properties of traditional anisotropic diffusion in the absence of small spatial features, and incorporating morphological operators into the diffusion coefficient simply necessitates that the larger gradient should not be translated to the position occupied by the smaller step function. Since the only spatial location guaranteed not to contain a smaller gradient is at the original location of the larger gradient, the simple criteria that must be satisfied by a morphological filter, which allows an anisotropic diffusion expression to smooth regions of low contrast, is that the morphological sequence produce a net translation of zero.

This condition on the construction of morphological filters that, when incorporated into the diffusion coefficient, allows the smoothing of regions of low contrast completes this section. Analysis of finite width edge models developed criteria on the morphological operators for the identification of small scale objects and required specific regions through which gradients must travel. Analysis of the second edge sequence introduced no further constraints on the path of gradient movement, but only defined the morphological filters to have a net gradient translation of zero.

63

Morphological operators, unlike the linear filters, exist which are capable of simultaneously satisfying these conditions. As an example, the morphological open-close filter is constructed from an erosion-dilation-dilation-erosion sequence and propagates the gradients throughout the necessary region of influence while introducing a net translation of zero. A visual example of incorporating this morphological sequence into the anisotropic diffusion coefficient is shown in Figure 4.10.

Study of the edge models has shown the success of the morphological filters in equipping the anisotropic diffusion equation with the capability of smoothing small scale objects without sacrificing its traditional ability to preserve edges. In the derivation of design criteria for the scale inducing morphological operators, only two edge sequences were considered, and though the step models are known to be eigenfunctions of the morphological system, the importance of the morphological anisotropic diffusion expressions in the following chapters motivates an experimental verification of the new morphological anisotropic diffusion process. The next section applies this new diffusion expression to the task of edge detection and presents performance results comparing traditional anisotropic diffusion and the enhanced linear scale aware diffusion representations.

**Figure 4.10** A visual example of filtering with the morphological anisotropic diffusion equation. The original image is located at the top, and its filtered result is located below. Notice how the diffusion process smoothes within the boundaries of an object, preserving edge locations, but is still capable of removing objects of small spatial scale. For comparison to the other diffusion expressions, refer to Figures 2.4 and 2.7.

Results

To display the effectiveness of the new morphological diffusion coefficient, simulations were conducted with three classes of the anisotropic diffusion coefficient. The first class was dependent solely on local gradient information and represented by the traditional diffusion coefficient expression, given as

$$c(\nabla \mathbf{I}) = e^{-\left(\frac{\|\nabla \mathbf{I}\|}{k}\right)^2}, \tag{4.6}$$

where $\mathbf{I}$ is the original image and $k$ is the gradient threshold. The second and third classes employed scale-modified definitions for the gradient. The second class used a linear filter in performing the gradient calculation, while the third class used a nonlinear morphological filter. These classes were represented by diffusion coefficients described by

$$c(\nabla \mathbf{I}) = e^{-\left(\frac{\|\nabla(\mathbf{G}_\sigma \ast \mathbf{I})\|}{k}\right)^2}, \tag{4.7}$$

where $\mathbf{G}_\sigma$ is a Gaussian kernel with standard deviation $\sigma$, and

$$c(\nabla \mathbf{I}) = e^{-\left(\frac{\|\nabla((\mathbf{I} \circ \mathbf{M}) \bullet \mathbf{M})\|}{k}\right)^2}, \tag{4.8}$$

where $\mathbf{M}$ is a morphological structuring element and $(\mathbf{I} \circ \mathbf{M}) \bullet \mathbf{M}$ is the image $\mathbf{I}$ filtered with an open-close filter.

Equivalent scale parameters for the linear and morphological filters in the scale aware diffusion coefficients were chosen to provide information removal of similar scale,

66

and both were defined by satisfying conditions necessary for subsampling a filtered representations by a factor of three. The Gaussian kernel used in the second coefficient class was defined to have a standard deviation of $6/\pi$, as suggested to approximately satisfy Shannon's sampling theorem in [5]. Similarly, the morphological kernel used in the third coefficient class was defined to be a square structuring element of width five, as suggested to satisfy the Homotopy Preserving Critical Sampling Theorem in [16].

Producing a qualitative evaluation of the three processes, the anisotropic diffusion equation was applied to synthetic imagery corrupted by 40% salt and pepper noise.[‡] Results are shown in Figure 4.11 (page 71). As can be seen, the diffusion equation based solely on local gradient information is unable to remove impulsive noise, while both spatially enlarged coefficients are capable of smoothing these small, high contrast objects and maintaining large scale edges. Results for coefficient classes two and three are visually similar, although closer inspection will show that the third class, the nonlinear morphological method, provides a slight improvement in feature definition.

A second qualitative example of the three anisotropic diffusion processes was attained by applying the smoothing operations to the cameraman image. These results are similar to those achieved with the previous synthetic imagery, and they are presented in Figure 4.12 (page 72). Again, the first coefficient class, using a traditional gradient calculation, is unable to remove fine detail, as evident by the existence of the small objects present on the ground. The second coefficient class, using a linear filter within its gradient calculation, removes these small features, but at the expense of introducing edge

---

[‡] All noise processes are defined to be white and uncorrelated.

67

movement and feature drift. (Notice the excessive smoothing of the camera and the movement of the elbow.) The new morphological anisotropic diffusion algorithm is capable of overcoming both deficiencies, removing small objects while maintaining edge locality.

While qualitative comparison of the three methods of anisotropic diffusion begins to distinguish the smoothing properties of the morphological diffusion coefficient, a quantitative comparison of their edge detection accuracy displays the advantages of the new diffusion expression. In determining the edge detection capabilities of the three variants of anisotropic diffusion, synthetic imagery was again corrupted by 40% salt and pepper noise. These images were then smoothed; and at each solution time, edges were identified and compared with known edge locations. Recognizing edges in the filtered imagery was accomplished with the use of a simple gradient based edge detector, well motivated by the smoothing properties of the anisotropic diffusion equation, and the threshold of the edge detector was defined to be equal to the gradient threshold of the diffusion coefficient, $k$.

Experimental comparison of edge detection performance was calculated using two quantitative metrics. The first, Pratt's edge quality measurement, is defined as

$$F = \frac{\sum\limits_{i=1}^{I_A} \dfrac{1}{1 + \alpha\left(\mathrm{d}(i)^2\right)}}{\max\{I_A, I_I\}} \tag{4.9}$$

where $I_A$ is the number of edge points detected in the filtered image result, $I_I$ is the number of edge points existing in the original, noise free imagery, $\mathrm{d}(i)$ is the Euclidean distance between an edge location in the original image and the nearest detected edge,

and $\alpha$ is a scaling constant, set to the suggested value of 1/9 [31]. Perfect recovery of all edge information in the original image results in an edge quality measurement of one (F=1); poor edge localization lowers the value.

The second measurement group contains two more tangible representations of the candidate filter performance, and the first measurement is defined to be the percentage of original edge points successfully identified by the edge detection process. Correctly recovering all edges in the initial image results in a 100% identification percentage, not detecting a feature at its original location lowers the identification measurement. The second measurement describes the ability of the edge detector to identify edges without detecting false edge locations. Expanding on the previous measurement, edge features which are not recognized and image locations which are erroneously classified as features are calculated. Perfect recovery of the original image results in an identification measurement of 100%, incorrect identification of any image location lowers the measurement.

The results of the numerical experiment are presented in Figure 4.13 (page 73). It may be seen that the linear coefficient initially outperforms the other diffusion variants in the edge quality measurement, but produces the poorest identification percentage. As solution time increases, the introduction of edge localization errors by the linear filter becomes more evident and is displayed by the rapid decrease in matched features. Specifically, at solution time three, the linear coefficient is unable to correctly identify the location of a single edge. The morphological anisotropic diffusion method provides

significant performance improvement, able to identify over 70% of the original edges and attain a solution quality measurement of 0.95.

The purpose of this section was to explore anisotropic diffusion processes that incorporate nonlinear filters into the diffusion coefficient, and the goal of this chapter was to show that these new diffusion expressions are capable of smoothing features of small spatial size while simultaneously smoothing regions of low contrast. Presentation of numerical simulations concludes this chapter and, coupled with the theoretical development of the previous sections, has presented a new morphological anisotropic diffusion coefficient, capable of removing fine features and noise while maintaining feature locality and reducing edge movement.

Development of a new diffusion expression is not the major goal of this thesis, however, and it serves only as a necessary step in the evolution of the anisotropic diffusion pyramid. In the next chapter, these morphological diffusion coefficients will be used to define scale parameters of the diffusion process. The parameter values will allow the anisotropic diffusion expression to produce signal representations suitable for sampling, and within the context of image pyramids, these scale depictions will correspond to individual levels of a pyramid structure. Definition of these diffusion parameters will result is the first formal definition of an anisotropic diffusion pyramid - the major goal of this thesis.
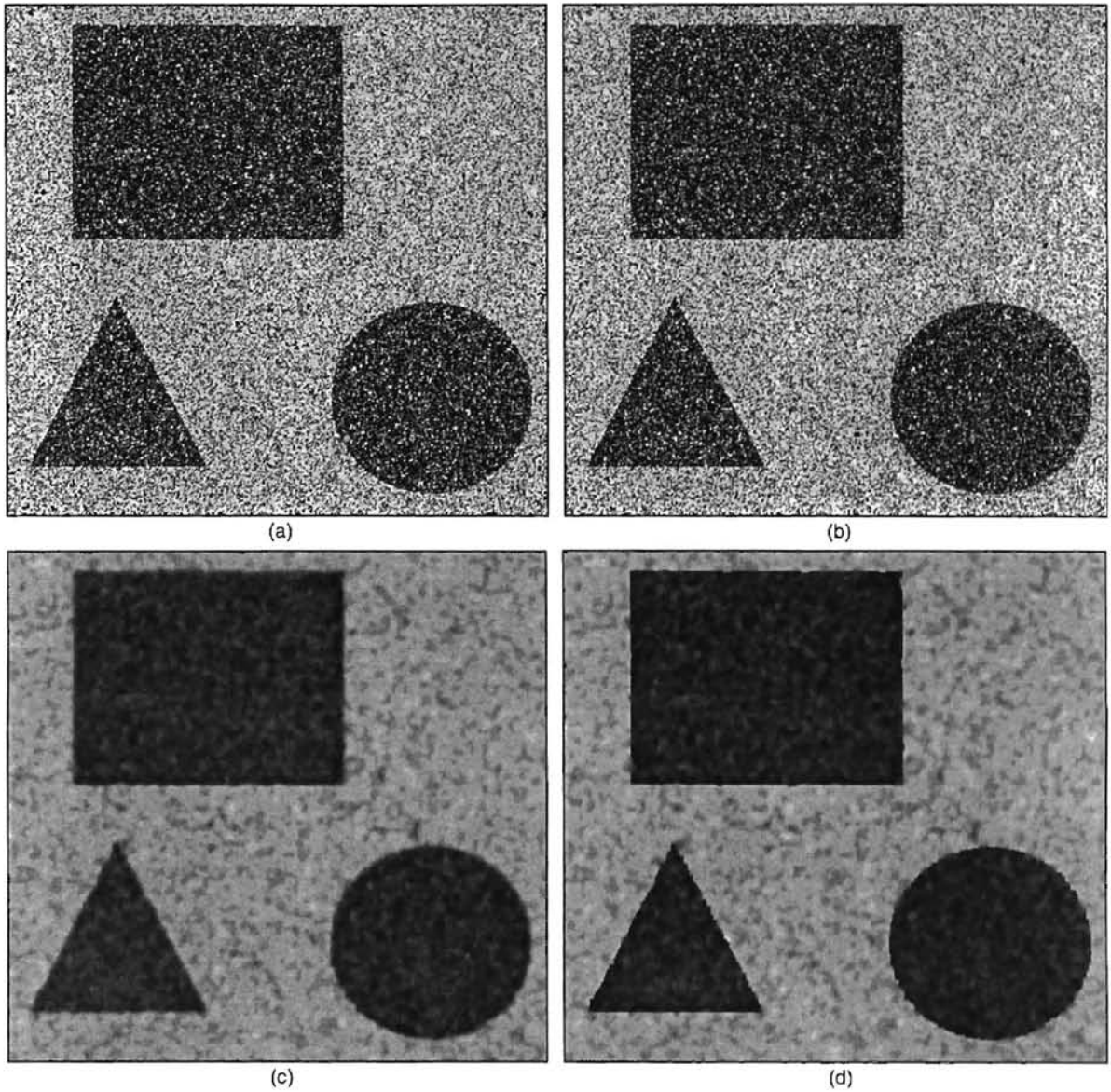
**Figure 4.11** Three classes of anisotropic diffusion applied to synthetic imagery: (a) original image corrupted with 40% salt and pepper noise; (b) results obtained using original anisotropic diffusion; (c) results obtained using traditional scale aware anisotropic diffusion; (d) results obtained using morphological anisotropic diffusion (4.8). The gradient threshold, $k=40$, and the solution time, $t=3$.

71

**Figure 4.12** Three classes of anisotropic diffusion applied to the cameraman image: (a) original image; (b) results obtained using original anisotropic diffusion; (c) results obtained using traditional scale aware anisotropic diffusion; (d) results obtained using morphological anisotropic diffusion (4.8). The gradient threshold, $k$=10, and the solution time, $t$=20.
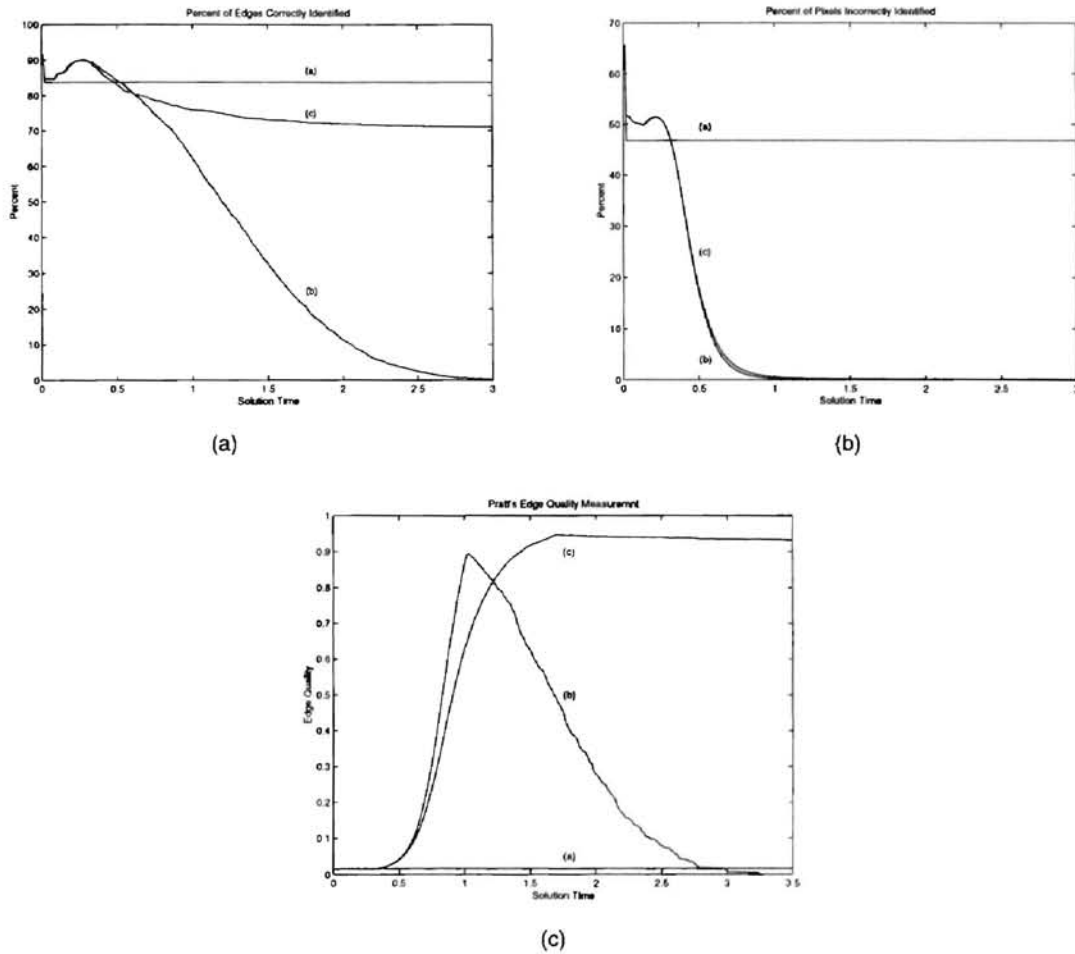
**Figure 4.13** Three quantitative measurements of edge detection performance: (a) percent of edges correctly identified; (b) percent of image pixels incorrectly classified by the edge detector; (c) results of Pratt's edge quality measurement.

73

# CHAPTER V

## THE ANISOTROPIC DIFFUSION PYRAMID

### Overview

Characterizing the smoothing performance of different anisotropic diffusion methods has been the major emphasis of the previous chapters. In the discussion, traditional representations of anisotropic diffusion have been shown to be ill-suited for prefiltering a signal before sampling, and original scale aware diffusion expressions, incorporating linear filters into their diffusion coefficients, have been shown to be incapable of simultaneously smoothing low contrast regions and smoothing small scale objects. In creating a smoothing mechanism capable of minimizing feature drift while providing necessary information removal, the last chapter has introduced nonlinear morphological operators into the diffusion coefficient. These diffusive processes have been shown to guarantee a reduction of information while preserving candidate edge locations.

Developing a scale aware anisotropic diffusion process with such ideal spatial smoothing characteristics is a major milestone of this thesis. However, it is not the primary objective. The goal of this thesis is to formally construct an anisotropic diffusion

pyramid and apply it to the problem of object identification and tracking. This is motivated since, theoretically, these multi-scale structures will contain a set of scale depictions which simplify establishing feature correspondence across scale and will allow the application of robust coarse to fine search procedures. Conceptually, the construction of an anisotropic diffusion pyramid will provide a very efficient and robust solution to the object identification problem.

The purpose of this chapter is to define the quantization of scale spaces generated by the anisotropic diffusion expression. Encapsulating scale space within an image pyramid requires the identification and selection of scale representations which satisfy sampling criteria, and in the construction of an anisotropic diffusion pyramid, this necessitates selecting filter parameters which produce a smoothed result suitable for sampling. Scale aware anisotropic diffusion expressions have multiple parameters of scale, and the goal of this chapter is to define parameter values such that the resulting anisotropic diffusion process will satisfy necessary sampling conditions. These smoothing parameters will be defined using a simple model of the scale aware diffusion process, presented in this chapter.

This chapter is organized as follows: The second section presents an ideal model of the scale aware anisotropic diffusion process, and the third section uses this model to define filter parameters for the anisotropic diffusion equation. These parameters allow the decimation of coarser scene depictions while ensuring spatial causality. The fourth section concludes this chapter and realizes a major goal of this thesis, formally defining an anisotropic diffusion pyramid.

Smoothing properties of the anisotropic diffusion expression are defined by the realization of their diffusion coefficients, and these diffusion coefficients are traditionally designed to preserve regions of high contrast while smoothing areas of low gradient. Several implementations of the anisotropic diffusion coefficient have been suggested in the literature, and the purpose of this section is to present a simple model of these processes for use in the next section, where smoothing parameters of the anisotropic diffusion pyramid are defined. Development of this model will abstract analysis of the diffusion equation from specific coefficient descriptions.

Previous chapters have exerted significant effort in discussing scale aware diffusion processes, and a definition of the ideal scale aware model begins by recognizing these initial descriptions. The first smoothing representation presented in this thesis was of traditional anisotropic diffusion, where in its analysis, the smoothing operator was modeled with a single smoothing threshold. Use of the smoothing threshold allowed traditional anisotropic diffusion to be described as a nonlinear filter, capable of preserving regions of significant gradient. It also suggested the unsuitability of using a diffusion operation as a sampling prefilter, showing that filtering sequences composed completely of high contrast objects did not remove any signal content.

Extensions to the traditional diffusion expressions attempted to solve these perceived deficiencies and produced the second diffusion model of this thesis. These scale aware diffusion equations were still capable of preserving significant edges, like

traditional diffusion realizations, but were designed to also smooth regions of small spatial size. Describing an ideal diffusion coefficient which models these performance constraints, prototypical scale aware diffusion expressions of the previous chapters have coefficient values near zero when in the presence of edges and coefficient values near one when in a small scale region.

Expanding this initial spatially aware diffusion model into a more complete representation simply requires describing the smoothing performance of the diffusion operator when not in the presence of large gradients or small spatial regions. Ideally, the anisotropic diffusion process should smooth areas of low contrast while preserving candidate edge locations, and a simple realization of this smoothing performance is attained by allowing uniform smoothing in all regions of small gradients or small spatial size. Identifying small scale objects by prefiltering the gradient used in the coefficient calculation with an abstract filtering function G, the ideal scale aware diffusion coefficient is defined to be

$$c(\nabla \mathbf{I}) = u\big(\|G(\nabla \mathbf{I})\| - k\big), \tag{5.1}$$

where c is the diffusion coefficient, $\nabla \mathbf{I}$ is the image gradient, $G(\nabla \mathbf{I})$ is the image gradient with small spatial features removed, u is the step function, and $k$ is the gradient threshold of the diffusion system. A possible realization of $G(\nabla \mathbf{I})$ might utilize the morphological open-close filter, presented in the previous chapter.

By assigning coefficient values of one to all locations with low gradient or with small spatial size and assigning smoothing values of zero to larger gradients belonging to large scale features, the ideal scale aware diffusion operator presents a very simple

filtering process and a useful tool for diffusion analysis. Inhibited only from smoothing across the boundaries of spatially large features, the diffusion coefficient smoothes all other locations within these boundaries uniformly. This filtering mechanism may be visualized as convolving the interior regions of a signal with a linear kernel while preserving significant edges and, effectively, a piece-wise linear filtering process. With constant coefficient values, the anisotropic diffusion expressions produce isotropic (or Gaussian) smoothing within these larger regions, and the performance of the ideal scale aware diffusion operator approximates a Gaussian lowpass filter at all locations of the original signal, except significant region boundaries.

Scale aware diffusion coefficients equip the anisotropic diffusion process with specific smoothing goals, and while the use of the ideal diffusion coefficient may have negative effects on the practical utility of a diffusion process, it does provide a simple model of previously presented diffusion coefficients[†]. Its definition creates a diffusion process which approximates Gaussian scale space at all locations internal to significant edges, while preserving these boundaries, and in the next section, this smoothing characteristic will be used to develop smoothing criteria for the suitable construction of an anisotropic diffusion pyramid.

---

† Though the smoothing model is simple, it is not necessarily trivial. Other filtering processes have been successfully modeled as binary decisions: the ideal low pass filter, for example.

Parameter Selection

Construction of an anisotropic diffusion pyramid requires selecting the scale parameters of the anisotropic diffusion equation which will allow the filtered result to be sampled without loss of information. In applying a pyramid structure to the problem of object identification and edge detection, sampling is not viewed within the context of reconstruction, but rather within the context of spatial causality. Spatial causality describes a cause and effect relationship between scale representations, and in this section, the construction of image pyramids which utilize the anisotropic diffusion equation as the scale generating operator will be considered. Throughout this discussion, the anisotropic diffusion equation will be analyzed using continuous input signals and treated as a piece-wise linear operator. Approximations of the continuous diffusion expressions with discrete difference equations will also be considered.

The ideal model of scale aware anisotropic diffusion, presented in the previous section, provides the outline for this section. Capable of smoothing regions of low contrast and smoothing regions of small spatial size, parameters of the scale aware anisotropic diffusion process will be developed which ensure that these smoothed regions will not introduce extraneous edge features into subsampled results. This section will first define smoothing criteria for filtering regions of low contrast and will then develop conditions ensuring proper smoothing of regions of small spatial size. Since the smoothing characteristic of the ideal scale aware diffusion expressions contains spatial and gradient based smoothing which are independent, satisfying these two smoothing

conditions defines an ideal diffusive process, well suited for application as a sampling prefilter, and a scale generating operator for image pyramid construction.

Sampling Regions of Low Contrast

Constructing an image pyramid with a fixed scale filter is possible only if the scale generating function also produces a signal suitable for sampling. For the traditional notion of scale space, encapsulated by the Gaussian pyramid, these sampling conditions are satisfied by the application of classical thoughts on sampling theory, and in considering the construction of an anisotropic diffusion equation, which ensures spatial causality within regions of low contrast, the smoothing conditions of an anisotropic diffusion pyramid and a Gaussian pyramid are equivalent. A region of low contrast is defined to be absent of internal edge features and, therefore, spatial causality is maintained if aliasing within the region is minimized.

Using a frequency domain analysis, a signal can be reconstructed from its sampled representation only if the original data is first band limited, and this is accomplished by prefiltering the original signal, removing all frequencies not accommodated by the new sample domain, so that its frequency response is below half the sampling rate. In practice, for construction of a Gaussian pyramid, the limited frequency response is usually approximated by setting the standard deviation of the Gaussian filter, $\sigma$, equal to twice the sample width divided by $\pi$, $2S/\pi$ [5].

The scale aware anisotropic diffusion process approximates a Gaussian filter within regions of low contrast, and the standard deviation of this filter is described by the

80

solution time parameter, $t$, of the anisotropic diffusion equation. The relationship between the width of a Gaussian filter and the solution time of the continuous time diffusion equation is defined as [36]

$$t = \frac{\sigma^2}{2},$$

(5.2)

and using this equality (and a uniform 1 of $S$ sampling scheme), a solution time satisfying the sampling requirements for image pyramid construction can be determined. Substituting $2S/\pi$, for the filter width, $\sigma$, in (5.2) presents a sampling solution time, $t_{LC}$, which ensures the maintenance of spatial causality within regions of low contrast. The solution time is expressed as

$$t_{LC} = \frac{2S^2}{\pi^2},$$

(5.3)

where $S$ is the decimation factor between pyramid levels. For a discrete approximation of the continuous diffusion process, this solution time describes the number of iterations the discrete system must be applied to the original signal, and for a discrete diffusion system to suitably smooth a region of low contrast the number of iterations is defined to be $t_{LC} \cdot \Delta t$, where $\Delta t$ is the time step used in the discrete realization of the diffusion equation [26].

The expression of a stopping time which removes all frequencies not supported in the next resolution domain ensures that subsampling regions of low contrast will not introduce edge features in coarser representations. However, this stopping time does not completely satisfy the sampling requirements of the anisotropic diffusion pyramid, as small spatial objects of high contrast may require additional smoothing to assure that they

do not interject edge features into subsampled scale depictions. To minimize the aliasing effects of sampling small scale regions, the diffusion process must remove all objects which cannot be spatially supported in the next resolution representation, reducing their gradients below the gradient threshold. In the next subsection, filter parameters will be defined which guarantee the spatial causality requirements of an image pyramid within these regions of small spatial size.

## Sampling Regions of Small Spatial Size

The elimination of a small region should result in its assignment to a larger one, and the purpose of this section is to develop smoothing criteria which will remove small, high contrast objects from subsampled scale space representations. Analyzing the anisotropic diffusion smoothing process as a linear filter within larger regions, the diffusion equation must smooth all internal features, reducing their gradients below the gradient threshold and accounting for the effects of sampling on the gradient measurement. The purpose of this subsection is to define smoothing criteria which will ensure spatial causality throughout a sampling operation, and the subsection begins by describing solution time parameters for smoothing spatially small features. After defining smoothing conditions which remove small scale objects from subsampled representations, the subsection concludes by defining the spatial size of an object which must be defined as "small" to maintain spatial causality.

The gradient of a sampled signal representation is proportional to the gradient of the original representation by the sample factor $S$. Guaranteeing that small spatial

features will be removed in coarser scale depictions is assured by smoothing all internal features such that all gradients are less than $k/S$, where $k$ is the gradient threshold of the diffusion system. Derivation of a solution time which satisfies this requirement is developed through the inspection of a single finite width step function of small spatial size and height $\alpha$. This function and its smoothed result are shown in Figure 5.1.

Treating the diffusion equation as a piece-wise linear filter, the smoothing conditions necessary for the removal of all edge features which belong to the finite width step function are derived by modeling the step function as a sequence of delta functions. This is allowable since the ideal scale aware diffusion equation is effectively a Gaussian linear filter within these spatially small regions, and inspecting the effect of smoothing a single impulse with the spatially aware diffusion expression allows the definition of stopping parameters for the filtering of small spatial objects. Figure 5.2 displays a single impulse function and its filtered result. The filtered representation is visually described as of Gaussian shape, and it may be mathematically
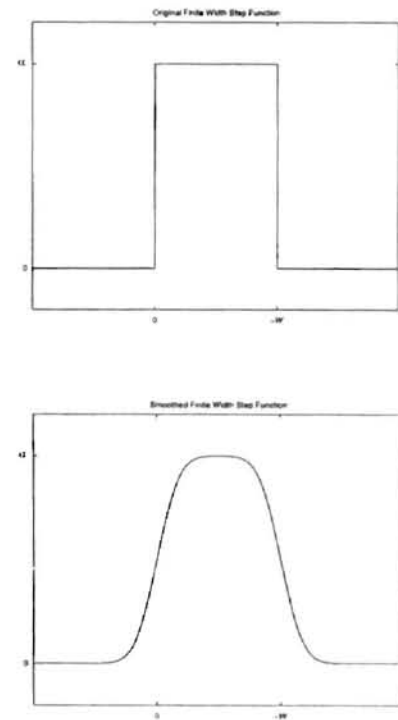


**Figure 5.1** The original finite width step function and an intermediate smoothed result. The filtered representation was created by smoothing the original function with an isotropic diffusion process (a Gaussian filter) for a fixed solution time.

expressed as

$$\alpha \cdot \left( \frac{1}{\sqrt{4\pi\ t}} e^{\frac{-x^2}{4t}} \right), \quad 0 < t < \infty \qquad (5.4)$$

where $\alpha$ is the magnitude of the impulse and $t$ is the solution time of the scale aware anisotropic diffusion expression.

Examining the derivative of a filtered impulse signal develops smoothing criteria guaranteeing that edge features are removed. Edges will be removed from subsampled representations if they are smoothed such that their gradients are less than $k/S$, and removal of a small finite width step function from coarser resolution representations ensures





**Figure 5.2** A single impulse function and its filtered result. Smoothing the impulse function with an isotropic diffusion process (a Gaussian filter) results in a filtered representation of Gaussian shape.

spatial causality within regions of small spatial features. Solution times assuring its removal from subsampled signal representations must satisfy

$$\left| \frac{\partial}{\partial x} \left( \alpha \cdot \frac{1}{\sqrt{4\pi\ t}} e^{\frac{-x^2}{4t}} \right) \right| < \frac{k}{S}, \quad 0 < t < \infty \qquad (5.5)$$

where $\alpha$ is the magnitude of the impulse, $k$ is the gradient threshold of the edge detection system, $t$ is the solution time of the scale aware anisotropic diffusion expression, and $S$ is the sample factor. Solving this equation, the minimum value of the spatial smoothing parameter, $t_{SS}$, is defined as

$$t_{ss} > \frac{\alpha \cdot S}{k\sqrt{8\pi}} \, e^{-\frac{1}{2}}, \tag{5.6}$$

which in the discrete approximation of the anisotropic diffusion equation is realized by smoothing the original signal with $t_{ss} \cdot \Delta t$ iterations of the diffusion expression, where $\Delta t$ is the time step used in the iterative realization.

Computation of an solution stopping time for actual construction of an anisotropic diffusion pyramid is accomplished by selecting the minimum value of the solution time parameter, $t$, which satisfies both the sampling requirements for smoothing regions of low contrast (5.3) and smoothing region of small spatial size (5.6). This is expressed as

$$t_s = \max\{t_{LC}, t_{ss}\}, \tag{5.7}$$

and for an iterative approximation of the anisotropic diffusion equation, the number of diffusion iterations necessary for filtering a signal such that it is suitable for sampling is approximately $t_s \cdot \Delta t$, where $\Delta t$ is the time step used in the discrete approximation of the diffusion equation.

This solution time parameter allows the scale aware anisotropic diffusion operation to smooth all regions of low contrast and all regions of small spatial size, assuring that edge features appearing in the subsampled coarse scale scene representations will not correspond to features in these smoothed regions. While this section has now defined the solution scale of the diffusion expression or, equivalently, the number of iterations of its discrete approximation, the ideal scale aware diffusion process contains a second description of scale. This additional filter parameter specifies the scope of the diffusion coefficient calculation and defines the spatial size of objects which are

determined to be "small" and subsequently removed. This parameter must be defined before the construction of image pyramids is possible.

Morphological operators have been used in the previous chapters to equip the scale aware anisotropic diffusion process with the capability of smoothing regions of low contrast and simultaneously smoothing regions of small spatial size. The operators are also important in defining the spatial size of objects which must be removed by the diffusion expression to ensure spatial causality. Borrowing from the mathematical study of topology, an image region may be viewed as a set and the region over which a coefficient is calculated defined. The diffusion system must preserve an object's *homotopy* across sampling domains, where homotopy is simply a one-to-one mapping of objects.

Similar to the frequency based sampling strategy, homotopy will only be guaranteed if all sets are removed which are spatially unsupported by the new sample domain. This requires the identification and smoothing of all objects smaller than the sample grid, so that after sampling they will not exist, and using morphological operators within a diffusion coefficient to identify these small scale objects, the scale of the diffusion coefficient may be defined by the structuring element size of the morphological operators. Ensuring the identification and removal of all objects smaller than the sample grid, the morphological operators must utilize structuring elements with diameter greater than the sample factor, and for constructing an anisotropic diffusion coefficient, the structuring elements used by the morphological filters must have diameter of $\sqrt{2} \cdot S$, where $S$ is the sample factor [16].

Definition of the scale of the morphological diffusion coefficient concludes this section, and will allow the forthcoming description of an anisotropic diffusion pyramid. The goal of this section was to define the two parameters of the diffusion operator such that spatial causality could be ensured throughout subsampled representations of anisotropic diffusion scale space. This was accomplished by considering the smoothing of regions with low contrast objects and regions of small spatial size and by developing conditions which assured that, after sampling, no extraneous edges would be introduced within these features. The development of smoothing parameters allows the anisotropic diffusion expression to be viewed as a suitable sampling prefilter and completes a major goal of this thesis. In the next section, a mathematical description of the construction of an anisotropic diffusion pyramid will be presented.

The Anisotropic Diffusion Pyramid

The goal of this thesis was to formally construct an image pyramid utilizing the anisotropic diffusion expression as the scale generating operator. In pursuing this goal, diffusion mechanisms have been developed which incorporate morphological operators into the diffusion coefficient and are capable of simultaneously smoothing regions of small spatial size and smoothing regions of low contrast. Maintaining spatial causality within an image pyramid is crucial for the application of traditional coarse to fine search procedures to discrete scale spaces, and development of diffusion parameters in this chapter ensures that spatial causality is maintained throughout a sampling operation. The

result is a smoothing operator capable of generating a scale space with minimal feature drift while satisfying conditions for prefiltering a signal before sampling. The purpose of this section is to formalize the anisotropic diffusion pyramid.

Image pyramid levels are constructed by successively filtering and subsampling previous resolution representations. Using the traditional discrete approximation of the partial differential equation presented in (2.3), construction of coarser resolution images within an anisotropic diffusion pyramid may be expressed as

$$\mathbf{I}_L = \left[ \mathbf{I}_{L-1} + \Delta t \sum_{m=1}^{t_s(\Delta t)} \left( c_N \nabla_N + c_S \nabla_S + c_E \nabla_E + c_W \nabla_W \right)_m \right]_{\downarrow S}, \qquad (5.8)$$

where $t_s$ is the solution time ensuring spatial causality from (5.7), $\downarrow S$ denotes subsampling by a factor of $S$, $\Delta t$ is the time step used in the numerical solution of the anisotropic diffusion expression, and $c$ is the ideal scale aware diffusion coefficient. For the generation of an anisotropic diffusion pyramid, the original image is first filtered with the anisotropic diffusion process with solution time parameters that assure the filtered result will be suitable for sampling. This intermediate filtered representation may then be subsampled, and this scale representation is referred to as the first pyramid level. (The original image is defined as the zero level within an image pyramid.) The higher levels of the multi-scale structure are computed by filtering and subsampling previous resolution representations.

Definitions of the solution time parameters in this chapter assume the use of an ideal scale aware diffusion coefficient. A visual example of an actual anisotropic diffusion pyramid is presented in Figure 5.3 and employs a more practical diffusion

damping coefficient. This scale aware diffusion coefficient, which is used in this section and the next chapter, is constructed by substituting a previously suggested realization of the scale aware diffusion coefficient [40], presented in (2.7), for the ideal scale aware diffusion expression. The diffusion coefficient used in the rest of this thesis is defined to be

$$
c = \begin{cases} 1, & \left\| \nabla((\mathbf{I} \bullet \mathbf{M}) \circ \mathbf{M}) \right\| < k \\ \left( \dfrac{k \sqrt{(k+1)}}{\left\| \nabla((\mathbf{I} \bullet \mathbf{M}) \circ \mathbf{M}) \right\| \sqrt{\left( \left\| \nabla((\mathbf{I} \bullet \mathbf{M}) \circ \mathbf{M}) \right\| + 1 \right)}} \right), & \left\| \nabla((\mathbf{I} \bullet \mathbf{M}) \circ \mathbf{M}) \right\| \geq k \end{cases}, \quad (5.9)
$$

where $k$ is the gradient threshold, $(\mathbf{I} \bullet \mathbf{M}) \circ \mathbf{M}$ denotes the close-open filtering of an image, and $\mathbf{M}$ is the structuring element, defined to have diameter greater than the sample factor, $\sqrt{2} \cdot S$.
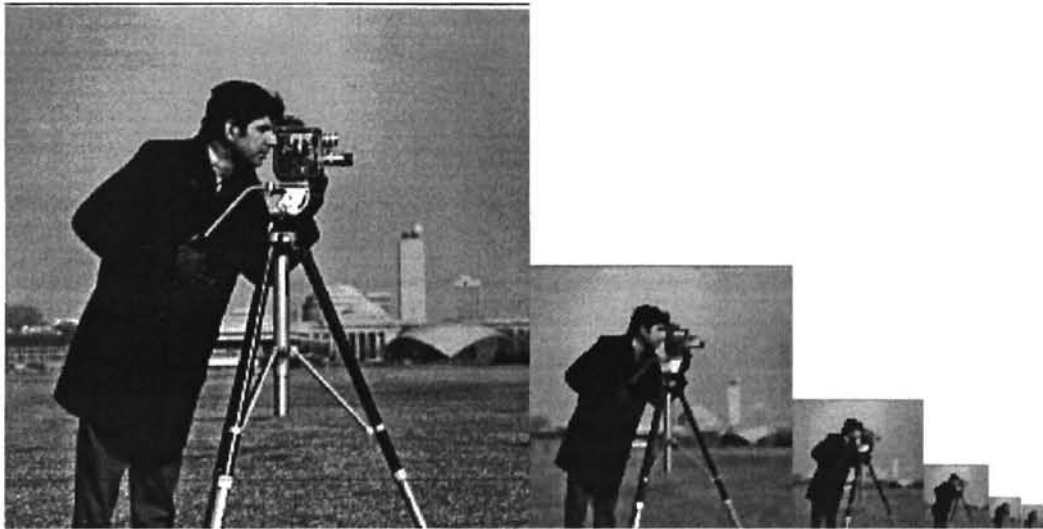


**Figure 5.3** Six levels of an anisotropic diffusion pyramid. The largest image is the original image. Smaller representations correspond to coarser scene depictions, each generated by filtering and subsampling the previous pyramid level. Notice that the cameraman is still visible in the coarsest level of the pyramid.

Presentations of the anisotropic diffusion pyramid concludes this chapter and realizes the major goal of this thesis. Its definition results in a scale space representation that preserves edge locations while providing necessary resolution reduction, and its creation is initially motivated by the theoretical promise of a robust and efficient solution to the object identification problem. In the next chapter, the integration of an anisotropic diffusion pyramid into an object identification and tracking system is explored. Results will be presented which show the computational efficiency and robustness of the new diffusion pyramid when applied to the identification problem. These results will also serve as verification of the pyramid construction techniques utilized in this chapter.

# CHAPTER VI

## ANISOTROPIC DIFFUSION PYRAMIDS:

## APPLICATION TO MULTI-SCALE TRACKING I

### Overview

The previous chapters have explored smoothing properties of several expressions of the anisotropic diffusion process. During this analysis, models of various diffusion characteristics were developed, and in the last chapter, these models were utilized in the formal construction of an anisotropic diffusion pyramid. The definition of an anisotropic diffusion pyramid presents a discrete scale space representation in which edge features are allowed minimal movement between scales, and theoretically, this should result in a very robust and efficient search method for an object identification problem. Creating a diffusion pyramid with these properties was the original motivation of this thesis.

The purpose of this chapter is to apply the anisotropic diffusion pyramid to a practical object identification problem, and the goal of this chapter is to measure the enhancement in solution quality provided by the pyramid structure. Displaying the improvements attained through the use of an anisotropic diffusion pyramid should indirectly verify its construction, and exhibiting the ability of the diffusion pyramid to

identify objects in a robust and efficient manner should validate the original motivations of this thesis. These results will complete the development of the anisotropic diffusion pyramid, showing that the newly created multi-scale structure does indeed present a robust and efficient solution to the object identification problem.

Solution quality improvements will be derived by incorporating the anisotropic diffusion pyramid into a target tracking system. As a prelude to assessing the quality of the tracker, the next section presents background on fundamental components of a multi-scale search. This search procedure initially locates an object at a coarse resolution representation within a pyramid, and it then utilizes these results to decrease the number of inspections in the finer scene representations. The performance gain provided by the coarse to fine search method is highly sensitive to the proper selection of the initial coarse resolution level, and the next section will also describe how these initial pyramid levels are defined.

The chapter is organized as follows: The second section provides background on the coarse to fine search procedure, discussing the realization of a multi-scale search within an image pyramid and the selection of the coarsest level in which the search procedure is initialized. The third section presents numerical simulations comparing traditional single resolution search techniques to methods incorporating the anisotropic diffusion pyramid. These results will show that the anisotropic diffusion pyramid provides a very robust and efficient solution to the object identification problem.

The advantage of using a multi-resolution search and track technique is embedded in the utilization of coarse scene representations for the initial identification of an object. Coarser scene information is created by successively filtering and subsampling the original image, and its use allows initial object identification to query information absent of noise and represented at reduced sample densities. Maximizing the benefits of these coarse to fine search procedures is accomplished by initially identifying an object at the coarsest resolution possible, and in a multi-scale search, this level is defined to be the *root level* of the search.

The purpose of this section is to discuss the application of coarse to fine search methods to the anisotropic diffusion pyramid, and in the next subsection the selection procedures necessary for identifying a root level will be presented. After developing criteria for determining the coarsest pyramid level used in a multi-scale search method, the second subsection will discuss the realization of a coarse to fine search within the anisotropic diffusion pyramid and its target tracking system.

Root Level Selection

The root level of an object simply defines the coarsest resolution representation in which the object will be identifiable. Consider the anisotropic diffusion pyramid presented in Figure 6.1, consisting of an infrared image of a plane in flight. Coarse to

fine search procedures will maximize both search efficiency and identification robustness by initially identifying this object at the coarsest resolution possible, and this level may be visually recognized as the third level of the pyramid structure. The purpose of this subsection is to describe the selection of this level for an arbitrary target object.



**Figure 6.1** An infrared image of a jet airplane in flight and its corresponding anisotropic diffusion pyramid. The coarsest representation which contains edge features belonging to the aircraft is the fourth largest image of the pyramid (the third level).

The first step in deriving root level definitions is to model the arbitrary object as a collection of smaller convex features. As an example, the jet aircraft in Figure 6.1 may be modeled as a composition of four smaller features: the fuselage, wing, tail, and landing gear. Expressing an arbitrary object, $O$, as the union of a set of smaller convex sets, the object may be described as

$$O = \bigcup_{m=1}^{M} o_m \quad where \ o_i \cap o_j = \varnothing \ for \ \forall i, j \le M, \tag{6.1}$$

where $O$ is the object of interest and $o_m$ are individual features. Considering all targets as a collection of convex features, the derivation of an object's root level is straightforward.

Defining the root level of an object with respect to its own internal composition is the first selection criteria considered in this subsection, and it necessitates defining the coarsest pyramid level that contains the object. The anisotropic diffusion expression removes all regions of small spatial scale, and for the complete removal of a target, the anisotropic diffusion process must remove all features of the target. These features will be removed according to tneir spatial size, with smaller features being removed before larger ones, and the selection of pyramid levels which contain the target require the representation to contain the largest target feature.

Subsampling a target feature reduces its spatial dimension, and in an anisotropic diffusion pyramid, spatial measurements of large objects, before and after sampling, are related by the proportionality factor $S$, where $S$ is the sample factor used for pyramid construction. For example, the original length of the aircraft fuselage in Figure 6.1 is 134 pixels. Measuring the length of the fuselage representation contained in the next pyramid level, constructed by filtering and subsampling the original image with a sample factor $S$, the sampled description of the fuselage has a length of 67 pixels. Therefore, subsampled objects have dimensions that are $S^{-1}$ the size of the original object, and the size of a large object at pyramid level $L$ may be expressed as

$$y_L = \frac{y}{S^L}, \tag{6.2}$$

95

where $y$ is the original measurement and $y_L$ is the equivalent spatial dimension in the subsampled domain.

Continuously filtering and subsampling an object should eventually result in its removal, and within an anisotropic diffusion pyramid, a feature is defined to be removed when its spatial size is smaller than the sample grid ($y_L < S$). The largest feature of the target will disappear in the construction of pyramid level $L+1$, when

$$S > \frac{y}{S^L},$$ (6.3)

where $y$ is the smallest spatial dimension of the feature (the minor axis), $S$ is the sample factor, and $L$ is the previous pyramid level.

Rearranging this equation produces the first definition of the root level of an object. The coarsest pyramid level in which a target will exist may be expressed as

$$L_R = \max\{L: L < \log_S |y| - 1\},$$ (6.4)

where $L_R$ is the root level defined by the internal characteristics of an object, $S$ is the sample factor used in pyramid construction, and $y$ is the minor axis of the largest convex feature of the target.

While the root level may be described by its internal composition, a complete definition must consider the content and construction of its environment. A target may also elude identification when multiple objects merge, as the search procedure will no longer be capable of resolving either individual object. Describing the separation distance between the target and the second object with the distance $z$, the two objects will merge in the construction of pyramid level $L+1$, when

$$S > \frac{z}{S^L}, \tag{6.5}$$

where $z$ is the minimum distance between the two objects, $S$ is the sample factor used in the construction of the pyramid, and $L$ is the previous pyramid level. Rearranging (6.5) presents the second description of the root level of a target, expressed as

$$L_R = \max\{L : L < \log_S |z| - 1\}, \tag{6.6}$$

where $L_R$ is the root level defined by the external characteristics of a scene.

The definition of the root level used in a coarse to fine search is determined both by an object's internal and external characteristics, and it may now be defined as

$$L_R = \max\{L : L < \log_S |d| - 1\}, \tag{6.7}$$

where

$$d = \min\{y, z\}. \tag{6.8}$$

The procedure used for selecting the root level of a coarse to fine search is best illustrated with an example. Consider the image presented in Figure 6.2, consisting of two aircraft. In determining the root level necessary for the identification of the smaller plane, two scene measurements must be



**Figure 6.2** Designing a coarse to fine search that is capable of identifying the smaller aircraft in the infrared image. Root level selection is dependent on two variables, the minor axis of the target's largest feature ($y$) and the minimum distance between the target and other objects of similar or greater scale ($z$).

97

determined. The first measurement is the minor axis length of the largest target feature, describing the pyramid level at which the target will be removed. In this example, the distance is denoted graphically with the variable $y$ and is measured to be 18 pixels. The second scene measurement necessary is the shortest distance between the smallest object and other large scale features. This measurement describes the pyramid level in which the two objects will merge, and in the figure, this distance is represented with the variable $z$ and measured to be 55 pixels. After identifying the two scene measurements, computation of the root level is accomplished by choosing the minimum of the two distance measurements and solving (6.7). In this example, the smallest scene measurement is the minor axis length of the target (18 pixels), and the root level of the smaller aircraft is defined to be the third level of the image pyramid, where the pyramid is constructed with a 1 of 2 sampling operation.

## Coarse to Fine Search Procedures

With the definition of the root level, a coarse to fine search procedure may be initialized to maximize the efficiency and structure of the anisotropic diffusion pyramid. These search methods identify an object within coarse resolution representations and then use these results to constrain higher resolution inspections. The practical realization of this search procedure is the topic of this subsection, and the goal is to describe the search method incorporated into the anisotropic diffusion pyramid and its target tracking system. This tracking system is used to present experimental measurements of the accuracy and
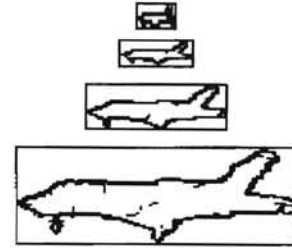
efficiency of the anisotropic diffusion pyramid in locating objects, and these results will be presented in the following section.

A coarse to fine search begins by identifying the target within the root level. In the target tracking system used in the next section, identification utilizes binary edge maps of the candidate target and the current scene. These edge maps are constructed by thresholding the gradient of the image, and an example of a multi-scale edge template is presented in Figure 6.3. Computing a binary exclusive-OR between coarse scale template information and the scene, and then summing the result, facilitates locating an object in the root level of the image pyramid. This operation may be described as



**Figure 6.3** A multi-resolution template for the anisotropic diffusion pyramid presented in Figure 6.1. Coarser template representations are used to search coarser scene descriptions within the pyramid.

$$Match(i, j) = \sum_x \sum_y \mathbf{T}_{L_R}(x, y) \oplus \mathbf{I}_{L_R}(x+i, y+j), \qquad (6.9)$$

where $\mathbf{T}_{L_R}$ is the template representation at the root level, $\mathbf{I}_{l_R}$ is the scene representation at the root level, and $\oplus$ denotes an exclusive-OR operation. Higher values for the binary template match correspond to higher similarity between template and scene, and in the simulation results to follow, the highest match is defined to correspond to the target.

Having found the best match between scene and template at location $(i,j)$ in level $L_R$ of the pyramid with the binary template match measurement, the goal of a coarse to fine search procedure is to use these results to guide and refine progressively higher resolution inspections. Anisotropic diffusion pyramids were designed to maintain spatial

causality, and this ensures that a features which exists at location $(i,j)$ in a coarse resolution representation will exist within the region $\left(S \cdot \left(i \pm \frac{1}{2}\right), S \cdot \left(j \pm \frac{1}{2}\right)\right)$ in higher resolution depictions, where $S$ is the sample factor used in the construction of the image pyramid [10]. Using this relationship between feature locations at different levels of a pyramid structure, the realization of a coarse to fine search procedure is straightforward. Target identification begins by locating the best match between edge template and scene, using (6.9). With these results, higher resolution information may be successively queried, and the binary template match needs only to be computed at four possible object locations. The identification results attained from inspecting the level $L_{R-1}$ are then used to constrain the search of the next level, $L_{R-2}$, and the procedure terminates after finding the target in the original image, $L_0$.

Presentation of the coarse to fine search method concludes necessary background on the construction of a target tracking system incorporating the anisotropic diffusion expression. In the next section, results will be presented which compare the solution quality of a traditional, single resolution, identification algorithm to the anisotropic diffusion tracking system. These results will show that the anisotropic diffusion pyramid does provide a robust and efficient solution to the object identification problem, verifying its construction and validating its creation.

Results

The purpose of this section is to present experimental results that attest to the solution quality of the anisotropic diffusion pyramid in object identification tasks. These results were attained by processing three "real world" image sequences with two object identification methods. The first method utilized the anisotropic diffusion tracking system, whose key components were discussed above, and the second method consisted of a traditional, single resolution, template matching algorithm[†]. Comparisons were then drawn which described the solution quality improvement provided by the anisotropic diffusion pyramid over a traditional search system. In the following simulations, solution quality will be described with two metrics: the measurement error between an object's identified location and ground truth and the computational requirements of the search routine. The results will show that the anisotropic diffusion pyramid provides a more robust and efficient solution to the object identification problem than traditional, single resolution, approaches.

---

[†] A single resolution template matching algorithm is analogous to a multi-scale search, with the root level defined to be the original image.

## Jet Sequence

The first image sequence used in the solution quality simulations consisted of 25 frames of a jet airplane in flight. Each original image, and the base of the corresponding pyramid, had a resolution of 256x256 pixels, and all pixels were capable of representing 256 intensity levels. The anisotropic diffusion pyramids were constructed with a 1 of 2 uniform sampling scheme, a gradient threshold ($k$) of 15, and a $\Delta t$ of ¼. (Figure 6.4 shows a montage of the image sequence, while Figure 6.1 displayed the pyramid constructed for the first frame of the sequence.) Implementing a multi-resolution search for the identification of the jet aircraft required the definition of the root level, and within the entire sequence, the largest element of the aircraft was its fuselage. Using (6.7), the root level of the sequence was defined to be the third resolution representation above the original image. The multi-scale template used in the simulations was shown in Figure 6.3.

Applying the coarse to fine search techniques to the target identification problem and using the third level of the pyramid as the root level, object recognition tasks were performed using the binary, edge based, template matching routine. The result was a significant increase in computational



**Figure 6.4** Six frames of the jet sequence. Shown are frames 1, 5, 10, 15, 20 and 25.

102

efficiency between the single resolution and multi-scale techniques. A single resolution match required approximately 172 seconds per frame on a Sun Ultra 1/170, while the multi-resolution approach required approximately 8 seconds per frame, including the construction costs of the anisotropic diffusion pyramid. The effect was a system performance improvement of 21 times tradition single-resolution methods.

Besides providing computation efficiency, multi-resolution techniques also increase system robustness. Using the same binary template matching routine, pixel localization errors were computed for both single and multi-resolution trials. These results are summarized in Figure 6.5, where the localization errors are expressed as the Euclidean distance between identified object locations and ground truth. For the first 14 frames of the sequence, the algorithms produce similar measurements. In the final 11 images, the multi-scale search was able to locate the target while the single-resolution method was not. This displays the inability of the single-resolution method to accommodate slight changes between the later images and the
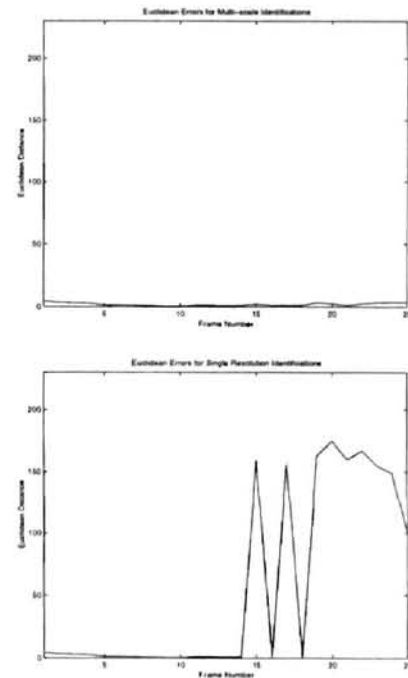


**Figure 6.5** Localization errors for the original jet sequence. Errors were calculated for both the single resolution and multi-resolution identification procedures by computing the Euclidean distance between identified target locations and known position information. The two algorithm produce equivalent results for the first 14 frames of the sequence. In the later frames of the sequence, the single resolution technique does not reliably identify the target.

template, while the anisotropic diffusion pyramids are more resilient, taking advantage of the high similarity between coarse scale descriptions within the image sequence.

To further display the robustness of the anisotropic diffusion pyramid, simulations were performed on the same sequence of images, but with each corrupted by Gaussian distributed noise. (The mean-square signal to noise ratio of the test images was 15.72, and Figure 6.6 shows the first frame of the sequence.) As can be seen from the identification results presented in Figure 6.7, the pixel localization error of the multi-scale technique increased in the presence of the additive noise, but the coarse to fine search method was still capable of providing acceptable estimates of the object location.



**Figure 6.6** The first frame of the jet sequence, corrupted with additive Gaussian noise.

Conversely, the single resolution method was unable to reliably determine the location of the target during any frame of the sequence. The ability to find an object in high clutter allowed the multi-scale object recognition system to provide a smaller pixel localization error, with a mean error of 3.69 pixels compared to 195.29 pixels of the single resolution system.
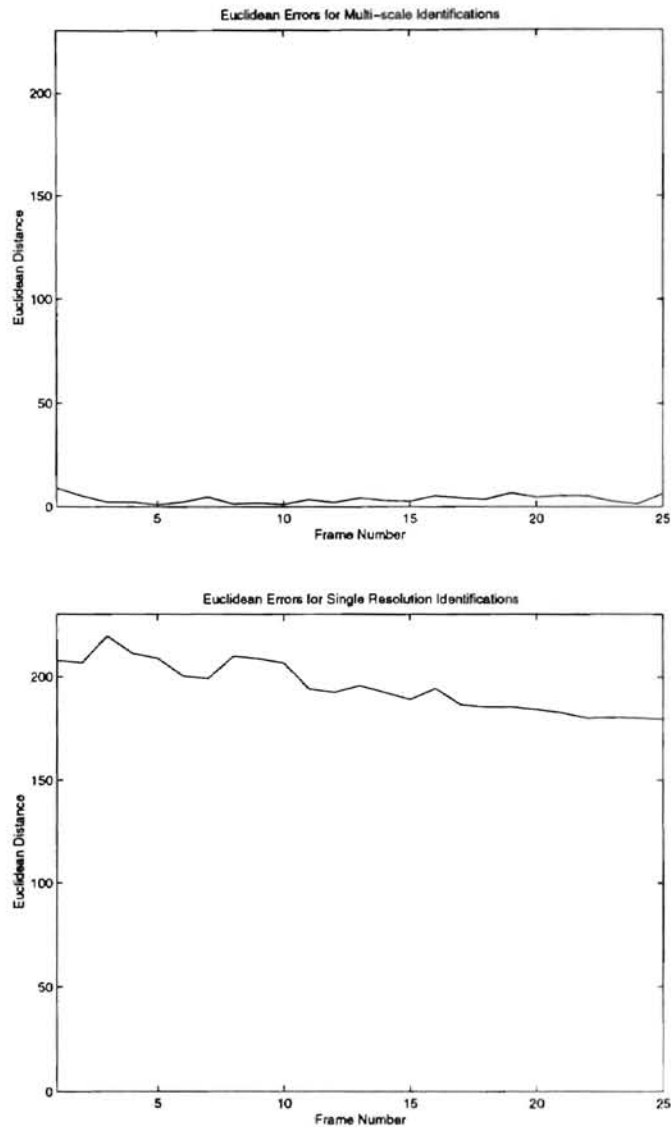
**Figure 6.7** Localization errors for the corrupted jet sequence. Errors were calculated for both the single resolution and multi-resolution identification procedures by computing the Euclidean distance between identified target locations and known position information. The anisotropic diffusion tracking system is capable of identifying the target in noisy imagery and only introduces small errors into the localization measurement. Single resolution identification techniques are not as robust, and the algorithm was not capable of correctly classifying the target in a single frame of the sequence.

## Semi Sequence

The second image sequence used for measuring the performance properties of the anisotropic diffusion tracking system consisted of 74 infrared images of a semi truck in motion. (Figure 6.8 shows an overview of the sequence.) All original images had a resolution of 320x240 pixels, and each pixel represents 256 intensity levels. The pyramids used for this evaluation were constructed with a 1 of 2 uniform sampling scheme, a gradient threshold ($k$) of 15, and a $\Delta t$ of $\frac{1}{4}$. (Figure 6.9 shows the pyramid constructed for the first frame of the sequence.) Implementing a multi-resolution search requires the root level of the target to be defined, and this necessitated the identification of the largest feature of the target. In the entire sequence, the most significant element of the semi was its trailer, and using (6.7), the root level of the sequence was defined to be the second resolution representation above the original image. Figure 6.10 shows the multi-scale edge template used in the simulations.

The first performance measurement of the simulation was the comparison of computational requirements between the multi-scale search and the single resolution identification procedure. Applying the multi-resolution techniques to the
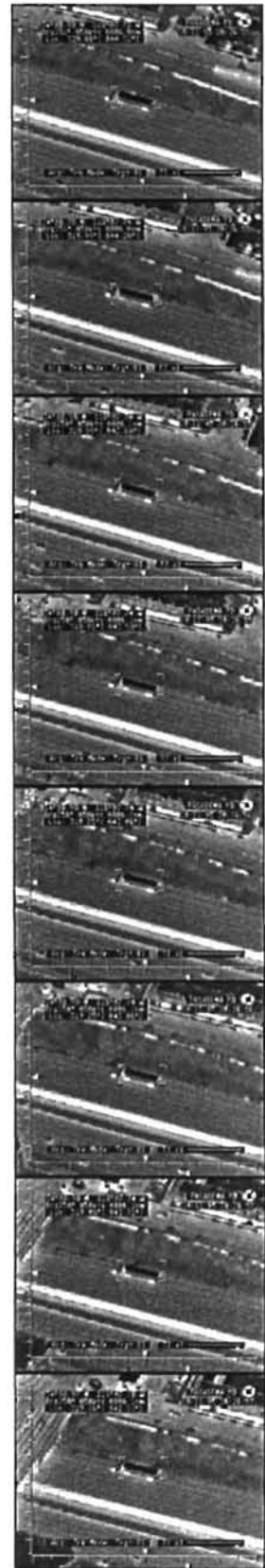


**Figure 6.8** Overview of the semi sequence.

106

target recognition problem and using the second level of the pyramid as the root level, the object identification tasks were performed using a binary template matching routine. For the single resolution match, the algorithm required approximately 46 seconds per frame on a Sun Ultra 1/170, while the multi-resolution technique needed approximately 6 seconds per frame. These results show an overall system performance improvement of 7.7 times tradition single-resolution methods.



**Figure 6.9** The first frame of the semi sequence and its corresponding anisotropic diffusion pyramid. The semi is visible only in the first three scene representations within the diffusion pyramid, resulting in the selection of the second level of the pyramid as the root level of the multi-scale search.



**Figure 6.10** The multi-scale template used for the semi sequence.

While the semi sequence again shows the presence of computational enhancements through the use of the anisotropic diffusion pyramid, the performance gains within this sequence account for only 1/3 of those attained with the previous jet aircraft simulation. As these two tracking sequences utilize different pyramid levels for their root level, the differences between the performance improvements within these images displays the sensitivity of the multi-scale method to root level selection. Coarser root levels allow more efficient and robust solutions then finer root levels.
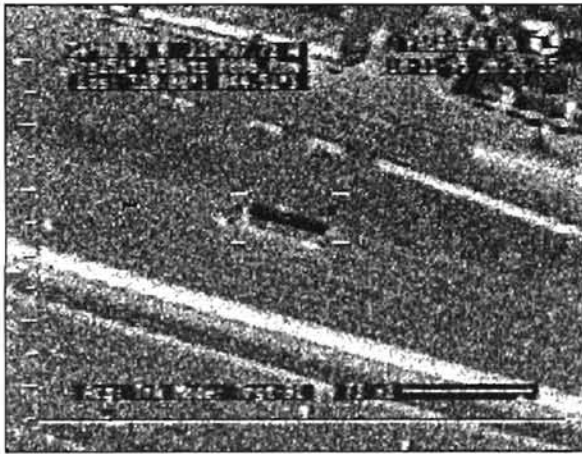
Using the same binary template matching routine, pixel localization errors were also computed for both single and multi-resolution trials. These results are summarized in Figure 6.11, where the localization errors are expressed as the Euclidean distance between the observed point and ground truth. These results show that the increased computational efficiency of the multi-scale search does not introduce extra localization error. For the entire sequence, the multi-scale and single-scale algorithms produce similar measurements. The mean localization error for the anisotropic diffusion tracking system was 1.08 pixels while the mean
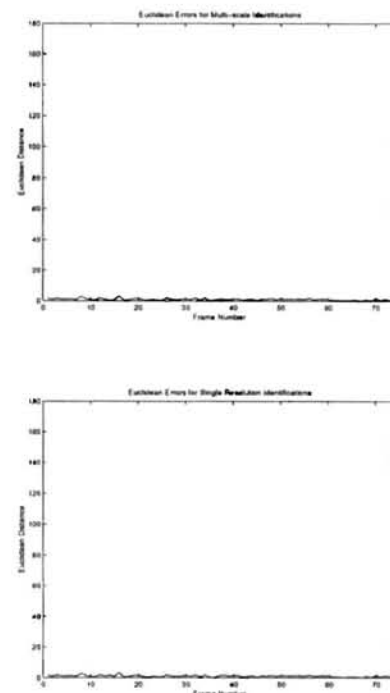


**Figure 6.11** Localization errors for the original semi sequence. Errors were calculated for both the single resolution and multi-resolution identification procedures by computing the Euclidean distance between identified target locations and known position information. The two algorithm produce comparable results, though the multi-scale technique requires less computational resources.

108

localization error for the single resolution technique was 1.11 pixels.

To display the robustness of the anisotropic diffusion identification system, the simulations were performed on the same set of images, but corrupted by Gaussian distributed noise. (The mean-square signal to noise ratio of the test images was approximately 15.34, and Figure 6.12 shows the first frame of the noisy sequence.) As can be seen from the identification results presented in Figure 6.13, the pixel localization error of the multi-scale technique increases in the presence of noise, but the algorithm is still capable of estimating the object location in the majority of the frames. The result of the single resolution method is very much in contrast, unable to locate the object during any frame of the corrupted sequence. The ability to find the target in noisy imagery allows the multi-scale object recognition system to provide a smaller pixel localization error, with a mean error of 19.15 pixels compared to 140.32 pixels of the single resolution system.



**Figure 6.12** The first frame of the semi sequence corrupted with Gaussian distributed noise.
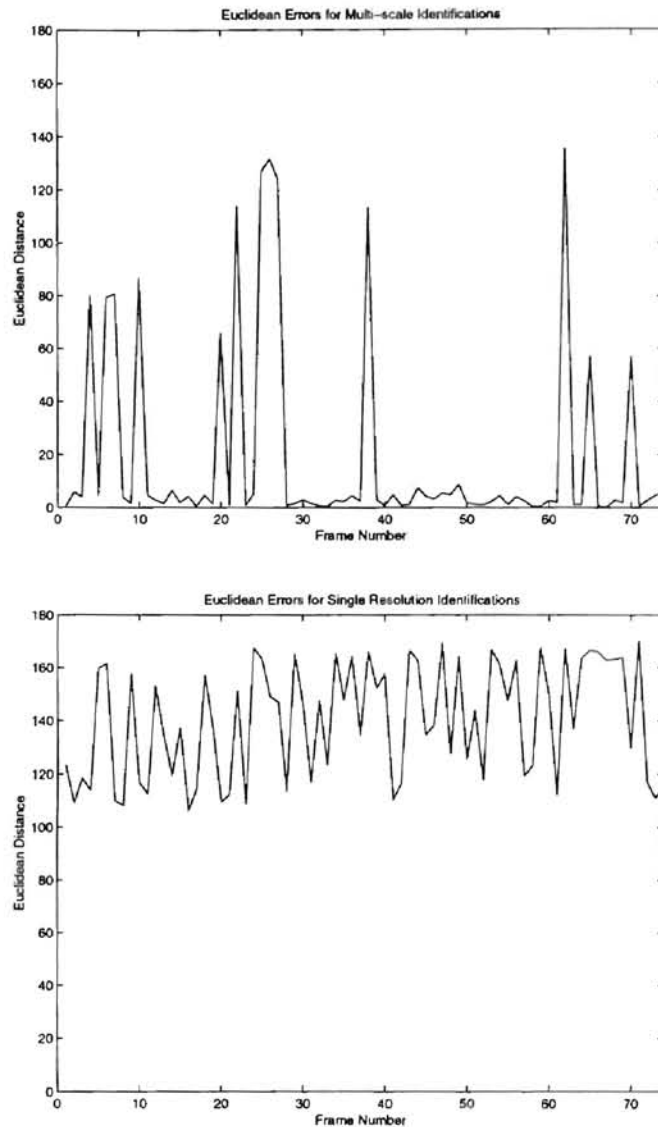
**Figure 6.13** Localization errors for the noisy semi sequence. Errors were calculated for both the single resolution and multi-resolution identification procedures by computing the Euclidean distance between identified target locations and known position information. The multi-scale algorithm is capable of classifying the target in a majority of the frames, denoted by the regions of low measurement error. Single resolution techniques are unable to find the target in any frame, denoted by large measurement error within each frame of the sequence.

## Truck Sequence

The final sequence used in the solution quality simulations consisted of 123 images of the rear of a truck. (Figure 6.14 shows several frames from the sequence.) The original images had a resolution of 320x240 pixels, and each pixel has a range of 256 intensity levels. The pyramids used for this evaluation were constructed with a 1 of 2 uniform sampling scheme, a gradient threshold ($k$) of 15, and a $\Delta t$ of ¼. (Figure 6.15 shows the pyramid constructed for the first frame.) To implement a multi-resolution search, the root level of the object must be identified, and in the entire sequence, the largest element of the truck is the back of its trailer. Using (6.5), the root level of the sequence is defined to be the fifth resolution representation above the original image. However, presence of other large objects in the scene necessitate selecting a lower initial level for the multi-scale search (in this example, the road and frame edge must be considered objects). In the following simulations, the root level of the target was defined to be the third resolution representation above the original image. Figure 6.16 shows the multi-scale edge template.
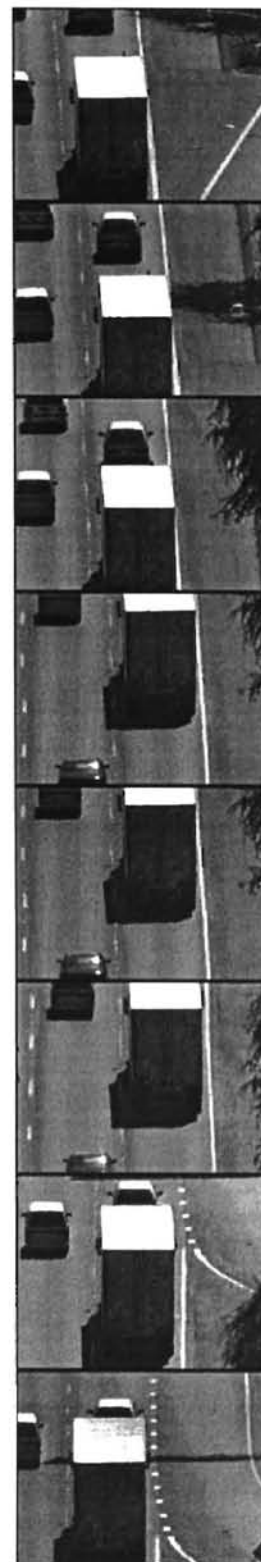


**Figure 6.14** Overview of the truck sequence.
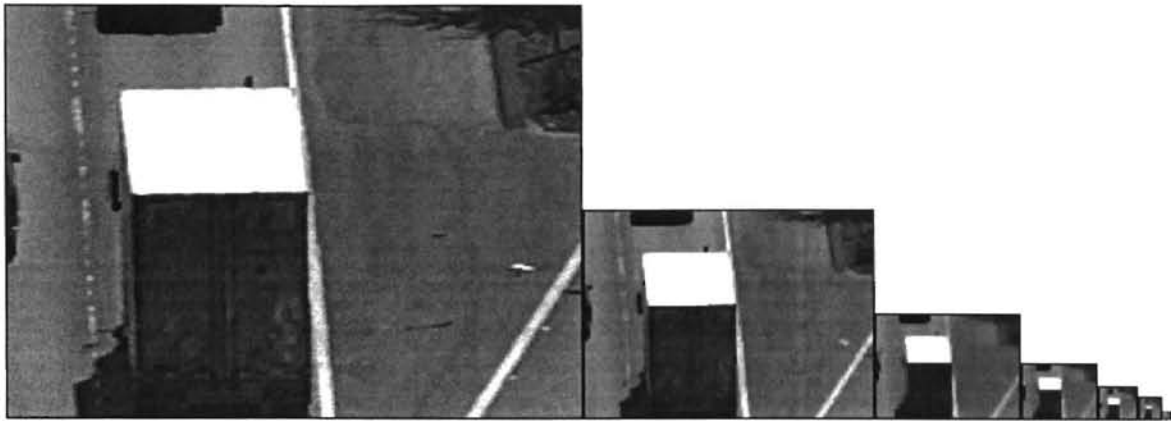
**Figure 6.15** The first frame of the truck sequence and its corresponding anisotropic diffusion pyramid. The truck is still visible in the sixth scene representation within the diffusion pyramid.
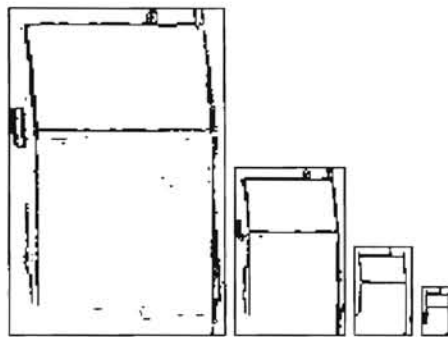


**Figure 6.16** The multi-scale template used for the truck sequence.

Applying multi-resolution techniques to the object recognition problem and using the third level of the pyramid as the root level, object recognition tasks were performed using a binary, edge based, template matching routine. For a single resolution match, the algorithm required approximately 325 seconds per frame on a Sun Ultra 1/170, while the multi-resolution technique required approximately 9 seconds per frame, including pyramid construction costs. The results show a system performance improvement of 36 times traditional single-resolution methods, again displaying the dependence of the anisotropic diffusion pyramid to the selection of the root level.

Increased computational efficiency does not introduce additional error into the identification results, and using the binary template matching routine, pixel localization errors were computed for both single and multi-resolution trials. These results are summarized in Figure 6.17, where the localization error is expressed as the Euclidean distance between the identified target location and ground truth. For the first 55 frames, the algorithms produce similar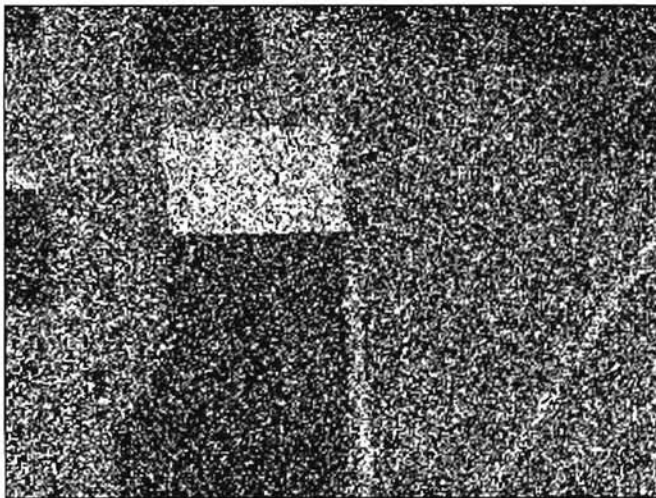 measurements. During the remaining images of the sequence, portions of the truck become occluded, with the top of the truck moving out of the image during frames 56 to 78 and the side of the truck occluded during the rest of the sequence. Both identification techniques are incapable of locating the target when the top of the truck is absent from the frame; however, the single-resolution method is also unable to accommodate the occlusion of the side panel in the later portions of the sequence. Anisotropic diffusion pyramids, and their coarse to fine search, are more resilient to these target changes, reacquiring the truck as it becomes entirely visible in the scene. Overall, the multi-scale
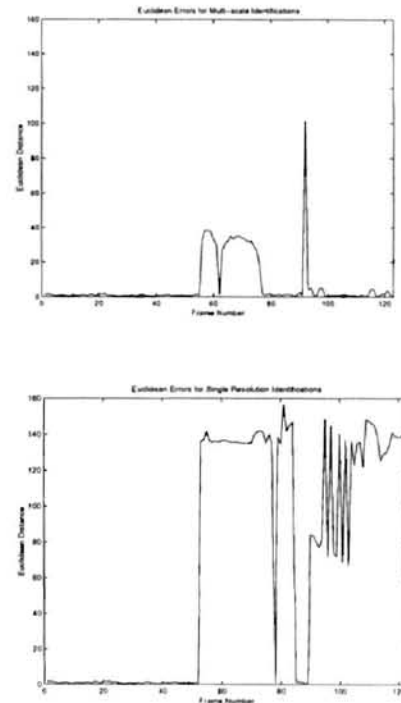


**Figure 6.17** Localization errors for the original truck sequence. Errors were calculated for both the single resolution and multi-resolution identification procedures by computing the Euclidean distance between identified target locations and known position information. The two algorithm initially produce comparable results. At approximately frame 55, significant portions of the truck become occluded. Upon reappearance, the multi-scale approach is capable of identifying the slightly deformed target while the single resolution technique is not.

113

technique had an average error of 7.06 pixels and the single resolution technique had an average error of 68.31 pixels.

Performing the simulations on a corrupted representation of the image sequence again displays the increased robustness of the anisotropic diffusion pyramid. (The image sequence was created by adding Gaussian noise to the original images, resulting in a mean-square signal to noise ratio of 2.64, and the first frame of the noisy sequence is show in Figure 6.18.) As can be seen from the data presented in Figure 6.19, the pixel localization error of the multi-scale technique increases in the presence of noise, while the single resolution method actually provides better results than attained on the original image set. The ability of the anisotropic diffusion pyramid to provide similar solutions to the identification problem in the presence of noise makes the multi-scale structure a more robust solution to the object identification problem and allows its mean error to increase by only 14.60 pixels. The mean error for the single resolution identification method decreased by 41.60 pixels, providing little correspondence to the original image sequence results.



**Figure 6.18** The first frame of the noisy truck sequence. The images were corrupted with Gaussian additive noise.
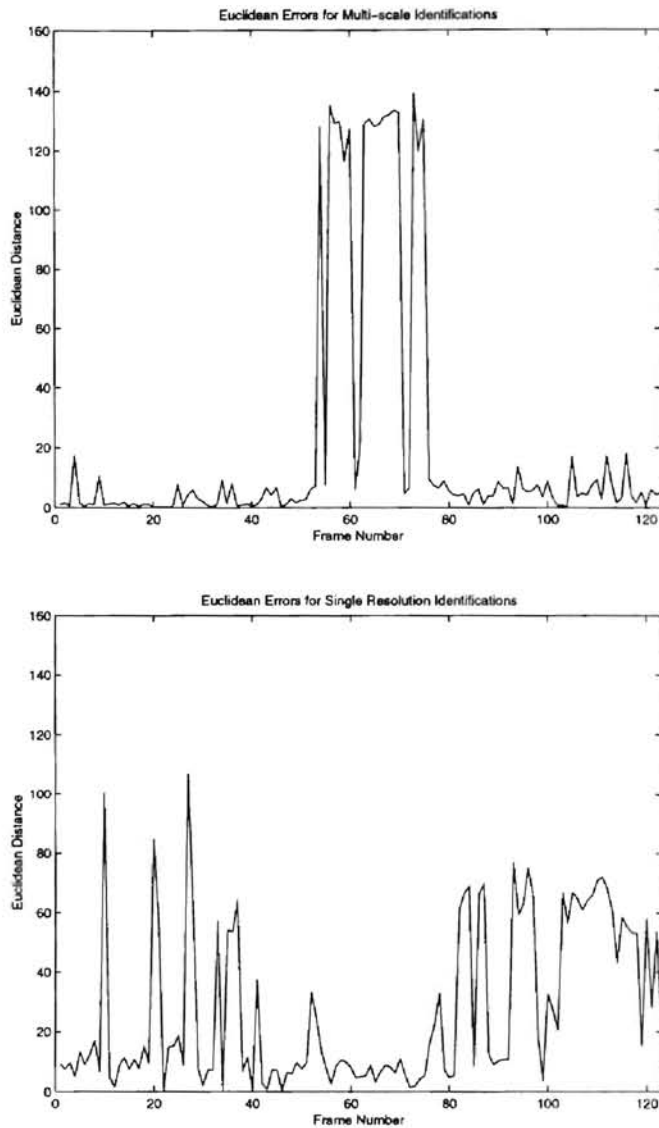
**Figure 6.19** Localization errors for the corrupted truck sequence. Errors were calculated for both the single resolution and multi-resolution identification procedures by computing the Euclidean distance between identified target locations and known position information. Using the anisotropic diffusion pyramid produces similar results between the original and noisy sequences. Application of traditional, single resolution techniques produce significant deviations in identification performance.

The purpose of this chapter was to apply the anisotropic diffusion pyramid to a practical object identification problem and to measure the introduced enhancement of solution quality. Reviewing the results presented for the three "real world" image sequences, it can be seen that the anisotropic diffusion pyramid does provide a robust and efficient solution to the object identification and tracking problem. Resulting in computational load improvements between 770-3600% over single resolution techniques (including the cost of image pyramid construction), the diffusion pyramid introduces significant processing reductions. This computational speed-up does not come at the expense of increased localization error in the identification of an object, as the multi-scale approach provides equivalent or lower errors than those attained with the single resolution methods. Anisotropic diffusion pyramids also proved to be robust, and the performance of the diffusion pyramids on the corrupted image sequences displays this solution characteristic. Using coarse scale information, absent of spurious noise and fine features, the diffusion pyramid was capable of identifying objects even in the presence of heavy noise. Single resolution methods were ineffective in these test sequences.

This chapter has shown that the diffusion pyramid does provide a robust and efficient solution to the object identification problem. Further solution quality improvements are possible, though, and they may be attained by incorporating temporal models of the target motion into the search procedure. A realistic anisotropic diffusion tracking system must consider the integration of the anisotropic diffusion process with predictive filters, as these tracking systems introduce additional refinements into the search procedure, allowing the system to filter measurement noise and provide location

116

estimates when a target becomes occluded. The formal construction of the anisotropic diffusion pyramid did not require these temporal models, nor did the experimental verification of the coarse to fine search procedure. However, development of an anisotropic diffusion tracking system would benefit from the inclusion of a predictive filter, and discussion of the integration of these temporal models with the anisotropic diffusion pyramid may be found in Appendix C.

# CHAPTER VII

## CONCLUSIONS AND FUTURE WORK

This thesis has formalized construction of an anisotropic diffusion pyramid. Toward this goal, a new morphological diffusion mechanism has been introduced and diffusion stopping time criteria has been developed. This criteria has allowed the nonlinear diffusion operation to serve as a suitable sampling prefilter, and completion of these tasks has allowed the generation of an image pyramid absent of edge movement. This thesis has quantized the continuous scale space of diffusion within an image pyramid, and for the anisotropic diffusion expression, completed the task which provided its original motivation. Construction of an anisotropic diffusion pyramid creates a discrete scale space representation and, coupled with a coarse to fine search, allows image features to be identified in a robust and efficient manor.

Introduction of the anisotropic diffusion pyramid has also allowed this thesis to consider the construction of a multi-scale target tracking system. This thesis has presented experimental results using three "real world' image sequences, and these results have shown the diffusion pyramid possesses robust identification properties, capable of locating a target in highly corrupted imagery. Experimental simulations have also displayed significant reductions in system computational requirements, as the multi-scale technique is able to identify objects approximately an order of magnitude faster than traditional, single resolution, procedures.

While this thesis has introduced a novel diffusion mechanism and an image pyramid with minimized feature movement, it has also presented many areas of future work. Scale aware anisotropic diffusion expressions, which incorporate linear filters into the diffusion coefficient, have been shown to be incapable of removing features of small spatial size while preserving edges. This property lead to the exploration of nonlinear filters within the diffusion coefficient and resulted in the presentation of a morphological anisotropic diffusion expression. While morphology is useful in the construction of an image pyramid, it is not suggested to be the "optimal" nonlinear filter for generating scale aware diffusion representations. Other nonlinear filters exist, and future work should consider these filters for diffusion coefficient integration.

Design of the anisotropic diffusion pyramid introduces another area with open tasks, as the smoothing parameters were derived using simple models of the diffusion expression. The experimental success of the resulting pyramid structure initially attests to the viability of this diffusion model, but analysis of actual diffusion mechanisms is needed to determine how they may deviate from this ideal representation. A second area of future work has also been presented in the construction of the anisotropic diffusion pyramid, as stopping time criteria were developed by considering the diffusion process in its continuous form and deriving solution parameters that allowed the removal of continuous features. Image processing applications operate on discrete imagery, and while the diffusion process was initially motivated by the continuous heat equation, its discrete realization is merely an approximation. Analysis of the difference equations used

in the application of the anisotropic diffusion expression may also reveal deviations between theoretical models and actual application.

Experimental results have shown the validity of integrating coarse to fine search methods into the anisotropic diffusion pyramid. By providing a scale space representation with minimal feature drift, the diffusion pyramid now provides a mechanism to explore multi-scale applications without the inherent correspondence problems of linear pyramids. Within the scope of target tracking applications, future work includes study of the coarse to fine search realization, automatic selection of root levels, integration of image pyramids with predictive filters, and maintenance of target templates. Other practical problems may also benefit from the use of the anisotropic diffusion pyramid, and these problems could include stereo matching, scene segmentation, image coding, and compression.

This thesis has been concerned with the identification of features within two-dimensional imagery. It has formalized the anisotropic diffusion pyramid, presented a novel diffusion coefficient expression, developed sampling conditions for the nonlinear diffusion process, and applied the new pyramid structure to a multi-scale target tracking system. Current acquisition technology demands robust and efficient solutions to the object identification problem. The anisotropic diffusion pyramid has been shown to provide these solutions.

# REFERENCES

[1] S.T. Acton, "Edge enhancement of infrared imagery by way of the anisotropic diffusion equation," in *Proceedings of the IEEE International Conference on Image Processing*, Lausanne, Switzerland, September 16-19, 1996.

[2] S.T. Acton, "Multigrid anisotropic diffusion," to appear in *IEEE Transactions on Image Processing*.

[3] J. Babaud, A.P. Witkin, M. Baudin and R. Duda, "Uniqueness of the gaussian kernel for scale-space filtering," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.8, no.1, pp.26-33, 1986.

[4] S. Biswas, N.R. Pa and S.K. Pal, "Smoothing of digital images using the concept of diffusion process," in *Pattern Recognition*, vol.29, no.3, pp. 497-510, 1996.

[5] P.J. Burt, "Fast filter transforms for image processing," in *Computer Graphics and Image Processing*, vol.16, no.1, pp.20-51, 1981.

[6] P.J. Burt, "Smart sensing within a pyramid vision machine," in *Proceedings of the IEEE*, vol.76, no.8, pp.1006-15, 1988.

[7] P.J. Burt, "Multiresolution techniques for image representation, analysis, and smart transmission," in *Proceedings of Visual Communications and Image Processing IV*, SPIE vol.1199, 1989.

[8] P.J. Burt and E.H. Adelson, "The Laplacian pyramid as a compact image code," in *IEEE Transactions on Communications*, vol.31, no.4, pp.532-40, 1983.

[9] P.J. Burt and E.H. Adelson, "A multiresolution spline with application to image mosaics," in *ACM Transactions on Graphics*, vol.2, no.4, pp.217-36, 1983.

[10] P.J. Burt, T. Hong and A. Rosenfeld, "Segmentation and estimation of region properties through cooperative hierarchical computation," in *IEEE Transactions on Systems, Man, and Cybernetics*, vol.11, no.12, pp.802-9 1981.

[11] V. Cantoni and S. Levialdi, "Multiprocessor Computing for Images," in *Proceedings of the IEEE*, vol.76, no.8, pp.959-69, 1988.

[12] L. Carrioli, "A pyramidal Haar-transform implementation", *Image Analysis and Processing*. New York: Plenum Publishing Corporation, 1986, pp.99-108.

[13] F. Catté, P.-L. Lions, J.-M. Morel and T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion," in *SIAM Journal of Numerical Analysis*, vol.29, no.1, pp.182-93, 1992.

[14] J.S Chen, "Generalized adaptive smoothing for multiscale edge detection," in *Applications of Artificial Intelligence X: Machine Vision and Robotics*, SPIE vol.1708, 1992.

[15] B. Fischl and E.L. Schwartz, "Adaptive nonlinear filtering for nonlinear anisotropic diffusion approximation in image processing," in *13th International Conference on Pattern Recognition*, Vienna, Austria, August. 25-30, 1996.

[16] D.A.F Florêncio and R.W. Schafer, "Homotopy and critical morphological sampling," in *Proceedings of the SPIE's 1994 International Symposium on Visual Communications and Image Processing*, Chicago, Illinios, September 1994.

[17] R.M. Haralick, S.R. Sternberg and X. Zhuang, "Image analysis using mathematical morphology," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.9, no.4, pp.532-50, 1987.

[18] J.P. Havlicek, D.S. Harding and A.C. Bovik, "The multicomponent AM-FM image representation," in *IEEE Transactions on Image Processing*, vol.5, no.6, pp.1094-100, 1996.

[19] H.J.A.M Heijmans and A. Toet, "Morphological sampling," in *CVGIP: Image Understanding*, vol. 54, pp. 384-400, 1991.

[20] R.E. Kalman, "A new approach to linear filtering and prediction problems," in *Journal of Basic Engineering*, pp.35-41, 1969.

[21] S. Kichenassamy, "The Perona-Malik paradox," to appear in *SIAM Journal of Numerical Analysis*.

[22] J.J. Koenderink, "The structure of images", in *Biological Cybernetics*, vol.50, pp.363-370, 1984.

[23] T. Lindeberg, "Scale-space for discrete signals" in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.12, no.3, pp. 234-54, 1990.

[24] P. Maragos, "Slope transforms: theory and application to nonlinear signal processing," in *IEEE Transactions on Signal Processing*, vol.43, no.4, pp.864-77, 1995.

[25]   D. Marr and E. Hildreth, "Theory of edge detection," in *Proceedings of the Royal Society of London*, vol.B207, pp.187-217, 1980.

[26]   W. Niessen, B.M. ter Haar Romeny and M. Viergever, "Numerical Analysis of Geometry-Driven Diffusion Equations", *Geometry-Driven Diffusion in Computer Vision*. Dordrecht: Kluwer Academic Publishers, 1994, ch.15, pp.393-411.

[27]   J.M. Ogden, E.H. Adelson, J.R. Bergen and P.J. Burt, "Pyramid-based computer graphics," in *RCA Engineer*, vol.30, pp.4-15, 1985.

[28]   C.F. Olsen and D.P. Huttenlocher, "Automatic target recognition by matching oriented edge pixels," in *IEEE Transactions on Image Processing*, vol.6, no.1, pp.103-13, 1997.

[29]   P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-12, no.7, pp.629-639, 1990.

[30]   P. Perona and M. Tartagni, "Diffusion networks for on-chip image contrast normalization," in *Proceedings of the IEEE International Conference on Image Processing*, Austin, Texas, November 13-16, 1996.

[31]   W.K. Pratt, *Digital Image Processing*. New York: Wiley, 1978, pp. 495-501.

[32]   P. Saint-Marc, J.S. Chen and G. Medioni, "Adaptive smoothing: a general tool for early vision" in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.13, no.6, pp. 514-29, 1991.

[33]   A. Toet, "A morphological image decomposition," in *Pattern Recognition Letters*, vol.9, pp.255-61, 1989.

[34]   F. Torkamani-Azar and K.E. Tait, "Image recovery using the anisotropic diffusion equation," in *IEEE Transactions on Image Processing*, vol.5, no.11, pp. 1573-78, 1996.

[35]   Y. Wang, J.S. Jin and J. Hiller, "An adaptive non-linear diffusion algorithm for image filtering," in *Proceedings of the SPIE: Real-Time Imaging II*, vol.3028, pp.26-37, 1997.

[36]   D.V. Widder, *The Heat Equation*. New York: Academic Press, 1975.

[37]   A.P. Witkin, "Scale-space filtering," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp.1019-22, 1983.

[38]  R.Y. Wong and E.L. Hall, "Sequential Hierarchical Scene Matching," in *IEEE Transactions on Computers*, vol.27, no.4, pp.359-66, 1978.

[39]  L. Wu and X. Xie, "Scaling theorems for zero-crossings" in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.12, no.1, pp. 46-54, 1990.

[40]  Y.-L. You, W. Xu, A. Tannenbaum and M. Kaveh, "Behavioral analysis of anisotropic diffusion in image processing," in *IEEE Transactions on Image Processing*, vol.5, no.11, pp. 1539-53, 1996.

[41]  A.L. Yuille and T.A. Poggio, "Scaling theorems for zero crossings" in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.8, no.1, pp. 15-25, 1996.

# APPENDIX

# APPENDIX A

## ANISOTROPIC DIFFUSION: A GRADIENT UPDATE EQUATION

One Dimension

Beginning with the one dimensional diffusion equation, the purpose of the first section is to present an equivalent diffusion expression based completely on signal gradient. The second section will extend these results to two dimensions. The original update equation is given as

$$\mathbf{I}_{t+\Delta t} = \mathbf{I}_t + \Delta t (c_E \nabla_E \mathbf{I} + c_W \nabla_W \mathbf{I}), \tag{A.1}$$

where $\mathbf{I}_{t+\Delta t}$ is the one dimensional signal at time $t+\Delta t$, $\mathbf{I}_t$ is the one dimensional signal at time value $t$, $\Delta t$ is the solution time step of the discrete system, $\nabla_W \mathbf{I}$ and $\nabla_E \mathbf{I}$ are the signal gradients in the east and west directions, respectively, and $c_E$ and $c_W$ are the diffusion coefficients in the east and west directions, respectively.

In the discrete equation, signal gradients and diffusion coefficients are calculated by simple differences. The east gradient, computed for signal location $x$, is expressed as

$$\nabla_E \mathbf{I}[x] = \mathbf{I}[x+1] - \mathbf{I}[x], \tag{A.2}$$

and the west gradient of the signal is expressed as

$$\nabla_W \mathbf{I}[x] = \mathbf{I}[x-1] - \mathbf{I}[x]. \tag{A.3}$$

126

Diffusion coefficients vary relative to these gradient magnitudes, and the computation of the two coefficients may be described as

$$c_E[x] = c\left(\left\|\nabla_E \mathbf{I}[x]\right\|\right) \tag{A.4}$$

and

$$c_W[x] = c\left(\left\|\nabla_W \mathbf{I}[x]\right\|\right). \tag{A.5}$$

Using the these signal gradient and diffusion coefficient expressions, the diffusion update equation may be simplified, described with a single gradient and diffusion coefficient. To accomplish this, define a new gradient and diffusion coefficient, where

$$\nabla \mathbf{I}[x] = \mathbf{I}[x-1] - \mathbf{I}[x] \tag{A.6}$$

and

$$c[x] = c\left(\left\|\nabla \mathbf{I}[x]\right\|\right). \tag{A.7}$$

Then, determine the relationship between the original gradients and coefficients.

$$\begin{aligned} \nabla \mathbf{I}_W[x] &= \mathbf{I}[x-1] - \mathbf{I}[x] \\ &= \nabla \mathbf{I}[x] \end{aligned} \tag{A.8}$$

$$\begin{aligned} \nabla \mathbf{I}_E[x] &= \mathbf{I}[x+1] - \mathbf{I}[x] \\ &= -(\mathbf{I}[x] - \mathbf{I}[x+1]) \\ &= -\nabla \mathbf{I}[x-1] \end{aligned} \tag{A.9}$$

$$\begin{aligned} c_W[x] &= c\left(\left\|\nabla_W \mathbf{I}[x]\right\|\right) \\ &= c\left(\left\|\nabla \mathbf{I}[x]\right\|\right) \\ &= c[x] \end{aligned} \tag{A.10}$$

$$c_E[x] = c(\|\nabla_E \mathbf{I}[x]\|)$$
$$= c(\|-\nabla \mathbf{I}[x-1]\|)$$
$$= c(\|\nabla \mathbf{I}[x-1]\|)$$
$$= c[x-1]$$

(A.11)

Finally, substitute the above equalities into (A.1)

$$\mathbf{I}_{t+\Delta t}[x] = \mathbf{I}_t[x] + \Delta t(-c[x-1]\nabla \mathbf{I}[x-1] + c[x]\nabla \mathbf{I}[x]).$$

(A.12)

Development of (A.12) allows the first goal of this chapter to be realized, developing an equivalent representation of the diffusion update equation based completely on signal gradient. Substituting (A.12) into (A.6) and simplifying, the expression for $\nabla \mathbf{I}_{t+\Delta t}[x]$ is found to be

$$\nabla \mathbf{I}_{t+\Delta t}[x] = \mathbf{I}_{t+\Delta t}[x+1] - \mathbf{I}_{t+\Delta t}[x]$$

(A.13)

$$\nabla \mathbf{I}_{t+\Delta t}[x] = \mathbf{I}_t[x+1] + \Delta t(-c[x]\nabla \mathbf{I}[x] + c[x+1]\nabla \mathbf{I}[x+1])$$
$$- \big(\mathbf{I}_t[x] + \Delta t(-c[x-1]\nabla \mathbf{I}[x-1] + c[x]\nabla \mathbf{I}[x])\big)$$

(A.14)

$$\nabla \mathbf{I}_{t+\Delta t}[x] = \nabla \mathbf{I}_t[x] + \Delta t(c[x+1]\nabla \mathbf{I}[x+1] - 2c[x]\nabla \mathbf{I}[x] + c[x-1]\nabla \mathbf{I}[x-1]). \quad \checkmark$$

Two Dimensions

Extension of the one dimensional update equation to two dimensional imagery is simplified by considering traditional discrete anisotropic diffusion realizations. The two dimensional update equation is expressed as

$$\mathbf{I}_{t+\Delta t} = \mathbf{I}_t + \Delta t(c_N \nabla_N \mathbf{I} + c_S \nabla_S \mathbf{I} + c_W \nabla_W \mathbf{I} + c_W \nabla_W \mathbf{I}),$$

(A.16)

where $\nabla_N I$ and $\nabla_S I$ are the image gradients in the north and south directions, respectively, and $c_N$ and $c_S$ are the diffusion coefficients in the north and south directions, respectively. This process may be separated into vertical and horizontal components, and Chapter 3 introduced the system:

$$\mathbf{I}_{t+\Delta t_H} = \mathbf{I}_{t+\Delta t_H} + \Delta t(c_N \nabla_N \mathbf{I} + c_S \nabla_S \mathbf{I}) \qquad (A.17)$$

and

$$\mathbf{I}_{t+\Delta t_H} = \mathbf{I}_t + \Delta t(c_W \nabla_W \mathbf{I} + c_W \nabla_W \mathbf{I}), \qquad (A.18)$$

where $\mathbf{I}_{t+\Delta t}$ is the signal after filtering in the horizontal direction.

The horizontal component of the two dimensional diffusion system is similar to the one dimensional equation considered in the previous section, except that the signal $\mathbf{I}$ possesses an additional dimension. Incorporating this second parameter into the difference expressions, gradient and coefficient parameters for the horizontal system component may be described as

$$\nabla_E \mathbf{I}[x, y] = \mathbf{I}[x+1, y] - \mathbf{I}[x, y], \qquad (A.19)$$

$$\nabla_W \mathbf{I}[x, y] = \mathbf{I}[x-1, y] - \mathbf{I}[x, y], \qquad (A.20)$$

$$c_E[x, y] = c(\|\nabla_E \mathbf{I}[x, y]\|), \qquad (A.21)$$

and

$$c_W[x, y] = c(\|\nabla_W \mathbf{I}[x, y]\|). \qquad (A.22)$$

The vertical component of the two dimensional diffusion system is calculated with respect to the second spatial variable $y$, with $x$ held constant. These parameters are defined as

129

$$\nabla_N \mathbf{I}[x, y] = \mathbf{I}[x, y+1] - \mathbf{I}[x, y], \qquad (A.23)$$

$$\nabla_S \mathbf{I}[x, y] = \mathbf{I}[x, y-1] - \mathbf{I}[x, y], \qquad (A.24)$$

$$c_N[x, y] = c\left(\left\|\nabla_N \mathbf{I}[x, y]\right\|\right), \qquad (A.25)$$

and

$$c_N[x, y] = c\left(\left\|\nabla_N \mathbf{I}[x, y]\right\|\right). \qquad (A.26)$$

Development of a two dimensional update expression parallels the work of the previous section, beginning with the definition of four new gradient and diffusion parameters. These parameters are defined as

$$\nabla \mathbf{I}[x] = \mathbf{I}[x-1, y] - \mathbf{I}[x, y] \qquad (A.27)$$

$$\nabla \mathbf{I}[y] = \mathbf{I}[x, y-1] - \mathbf{I}[x, y] \qquad (A.28)$$

$$c[x] = c\left(\left\|\nabla \mathbf{I}[x]\right\|\right). \qquad (A.29)$$

and

$$c[y] = c\left(\left\|\nabla \mathbf{I}[y]\right\|\right). \qquad (A.30)$$

Next, the relationship between the original diffusion parameters and the new gradient and coefficient expressions are determined. These equalities are described as:

$$\begin{aligned}
\nabla \mathbf{I}_N[x, y] &= \mathbf{I}[x, y+1] - \mathbf{I}[x, y] \\
&= -\left(\mathbf{I}[x, y] - \mathbf{I}[x, y+1]\right) \\
&= -\nabla \mathbf{I}[y-1]
\end{aligned} \qquad (A.31)$$

$$\begin{aligned}
\nabla \mathbf{I}_S[x] &= \mathbf{I}[x, y-1] - \mathbf{I}[x, y] \\
&= \nabla \mathbf{I}[y]
\end{aligned} \qquad (A.32)$$

$$\nabla \mathbf{I}_E[x, y] = \mathbf{I}[x+1, y] - \mathbf{I}[x, y]$$
$$= -(\mathbf{I}[x, y] - \mathbf{I}[x+1, y]) \tag{A.33}$$
$$= -\nabla \mathbf{I}[x-1]$$

$$\nabla \mathbf{I}_W[x, y] = \mathbf{I}[x-1, y] - \mathbf{I}[x, y]$$
$$= \nabla \mathbf{I}[x] \tag{A.34}$$

$$c_N[x, y] = c\big(\|\nabla_N \mathbf{I}[x, y]\|\big)$$
$$= c\big(\|-\nabla \mathbf{I}[y-1]\|\big)$$
$$= c\big(\|\nabla \mathbf{I}[y-1]\|\big) \tag{A.35}$$
$$= c[y-1]$$

$$c_S[x, y] = c\big(\|\nabla_S \mathbf{I}[x, y]\|\big)$$
$$= c\big(\|\nabla \mathbf{I}[y]\|\big) \tag{A.36}$$
$$= c[y]$$

$$c_E[x, y] = c\big(\|\nabla_E \mathbf{I}[x, y]\|\big)$$
$$= c\big(\|-\nabla \mathbf{I}[x-1]\|\big)$$
$$= c\big(\|\nabla \mathbf{I}[x-1]\|\big) \tag{A.35}$$
$$= c[x-1]$$

$$c_W[x, y] = c\big(\|\nabla_W \mathbf{I}[x, y]\|\big)$$
$$= c\big(\|\nabla \mathbf{I}[x]\|\big) \tag{A.36}$$
$$= c[x]$$

Substituting these relationships into (A.17) and (A.18) presents an intermediate expression of the two dimensional diffusion equation. This process is described with the vertical and horizontal components:

$$\mathbf{I}_{t+\Delta t}[x, y] = \mathbf{I}_{t+\Delta t_H}[x, y] + \Delta t(-c[y-1]\nabla \mathbf{I}[y-1] + c[y]\nabla \mathbf{I}[y]) \tag{A.37}$$

and

131

$$I_{t+\Delta t_H}[x,y] = I_t[x,y] + \Delta t(-c[x-1]\nabla I[x-1] + c[x]\nabla I[x]). \qquad (A.38)$$

Development of (A.37) and (A.38) allows the second goal of this chapter to be realized, extending the gradient based diffusion update equation to two dimensions. The expression for the horizontal component of the diffusion system is derived by substituting (A.38) into (A.27) and (A.28). The simplified expression becomes

$$\begin{aligned}\nabla I_{t+\Delta t_H}[x] &= I_{t+\Delta t_H}[x+1,y] - I_{t+\Delta t_H}[x,y]\\ \nabla I_{t+\Delta t_H}[y] &= I_{t+\Delta t_H}[x,y+1] - I_{t+\Delta t_H}[x,y]\end{aligned} \qquad (A.39)$$

$$\begin{aligned}\nabla I_{t+\Delta t_H}[x] &= I_t[x+1,y] + \Delta t(-c[x]\nabla I[x] + c[x+1]\nabla I[x+1])\\ &\quad - \Big(I_t[x,y] + \Delta t(-c[x-1]\nabla I[x-1] + c[x]\nabla I[x])\Big)\\ \nabla I_{t+\Delta t_H}[y] &= I_t[x,y+1] + \Delta t(-c[x]\nabla I[x] + c[x]\nabla I[x])\\ &\quad - \Big(I_t[x,y] + \Delta t(-c[x]\nabla I[x] + c[x]\nabla I[x])\Big)\end{aligned} \qquad (A.40)$$

$$\begin{aligned}\nabla I_{t+\Delta t_H}[x] &= \nabla I_t[x] + \Delta t(c[x+1]\nabla I[x+1] - 2c[x]\nabla I[x] + c[x-1]\nabla I[x-1])\\ \nabla I_{t+\Delta t_H}[y] &= \nabla I_t[y]\end{aligned} \qquad \checkmark$$

The gradient update expression for the vertical component of the diffusion system is derived by substituting (A.37) into (A.27) and (A.28). The simplified expression becomes

$$\begin{aligned}\nabla I_{t+\Delta t}[x] &= I_{t+\Delta t}[x+1,y] - I_{t+\Delta t}[x,y]\\ \nabla I_{t+\Delta t}[y] &= I_{t+\Delta t}[x,y+1] - I_{t+\Delta t}[x,y]\end{aligned} \qquad (A.41)$$

$$\begin{aligned}\nabla I_{t+\Delta t}[x] &= I_{t+\Delta t_H}[x+1,y] + \Delta t(-c[y]\nabla I[y] + c[y]\nabla I[y])\\ &\quad - \Big(I_{t+\Delta t_H}[x,y] + \Delta t(-c[y]\nabla I[y] + c[y]\nabla I[y])\Big)\\ \nabla I_{t+\Delta t}[y] &= I_{t+\Delta t_H}[x,y+1] + \Delta t(-c[y]\nabla I[y]] + c[y+1]\nabla I[y+1])\\ &\quad - \Big(I_{t+\Delta t_H}[x,y] + \Delta t(-c[y-1]\nabla I[y-1] + c[y]\nabla I[y])\Big)\end{aligned} \qquad (A.42)$$

132

$$\nabla \mathbf{I}_{t+\Delta t}[x] = \nabla \mathbf{I}_{t+\Delta t_H}[x]$$

$$\nabla \mathbf{I}_{t+\Delta t}[y] = \nabla \mathbf{I}_{t+\Delta t_H}[y] + \Delta t \big(c[y+1]\nabla \mathbf{I}[y+1] - 2c[y]\nabla \mathbf{I}[y] + c[y-1]\nabla \mathbf{I}[y-1]\big)$$

# APPENDIX B

## EDGE PRESERVATION IN MONOTONIC REGIONS

**Theorem**

*If there exists a function* $\beta(K_{DC},t)$ *such that for any* $|\alpha| \geq \beta(K_{DC},t)$, $c(\alpha) \cong 0$; *then the discrete anisotropic diffusion equation will preserve all regions with gradient magnitude greater than* $\alpha$, *if the gradient exists within a monotonic region.*

**Proof**

*Given a locally monotonic sequence,* S, *and a diffusion coefficient* c, *where* $c(x) \geq 0 \ \forall x$, *the diffusion update equation is expressed as*

$$\nabla S_{\Delta t}[x] = \nabla S[x] + \Delta t\big(c(\nabla S[x-1])\nabla S[x-1] - 2c(\nabla S[x])\nabla S[x] + c(\nabla S[x+1])\nabla S[x+1]\big).$$

*where* $\Delta t \geq 0$.

Now, let $\nabla S[x]$ be a gradient of significant magnitude. $\therefore |\nabla S[x]| \geq \beta(K_{DC},t)$ and $c(\nabla S[x]) \cong 0$. Incorporating this into the update equation produces a simplified form, expressed as $\nabla S_{\Delta t}[x] = \nabla S[x] + \Delta t\big(c(\nabla S[x-1])\nabla S[x-1] + c(\nabla S[x+1])\nabla S[x+1]\big)$.

Since **S** is defined to be monotonic in a neighborhood around $x$, the sign of the gradients in the update equation do not change sign. Recall that the diffusion coefficient is defined to be positive, $c(x) \geq 0 \ \forall x$, then the update will always have the same sign as $\nabla S[x]$, and the gradient magnitude of $\nabla S[x]$ will never diminish as

$$\big|\ \nabla S[x]\big| + \big|\nabla S[x-1]\big| + \big|\nabla S[x+1]\big| = \big|\ \nabla S[x] + \nabla S[x-1] + \nabla S[x+1]\ \big|.$$

Thus, $\big|\ \nabla S[x] + \Delta t\big(c(\nabla S[x-1])\nabla S[x-1] + c(\nabla S[x+1])\nabla S[x+1]\big)\big| \geq |\nabla S[x]|$, which can be simplified. Substituting $\nabla S_{\Delta t}[x]$ for the left hand side of the inequality, the equation reduces to $\big|\ \nabla S_{\Delta t}[x]\ \big| \geq \big|\nabla S[x]\big|$.

Therefore, if $|\nabla S[x]| \geq \beta(K_{DC},t)$ then $|\nabla S_{\Delta t}[x]| \geq \beta(K_{DC},t)$, $\forall \Delta t$. ✔

# APPENDIX C

# ANISOTROPIC DIFFUSION PYRAMIDS:

## APPLICATION TO MULTI-SCALE TRACKING II

### Overview

Locating an object using the anisotropic diffusion pyramid can provide significant improvement in system robustness and computation efficiency. Chapter 6 presented experimental results verifying this solution quality increase, and these results were generated by incorporating the anisotropic diffusion pyramid into a target tracking system and measuring the difference between locations identified as containing a target and ground truth. Showing the improved performance properties concluded development of the anisotropic diffusion pyramid and validated the original motivations for its construction.

Target tracking systems traditionally incorporate temporal models of the tracked object into their identification procedure, and this information provides further increases in computational efficiency and system robustness. Complete realization of a target tracking system utilizing and anisotropic diffusion pyramid requires the integration of predictive filters, and the purpose of this appendix is to discuss methods for incorporating diffusion pyramids into a target tracking system. The goal of this appendix is to introduce

135

three possible design strategies for the construction of a multi-scale target tracking system and to provide experimental results comparing these three methods on an actual tracking sequence.

This appendix is organized as follows: The second section presents necessary background on the Kalman filter. This filter provides an optimal linear mean squared error prediction of the target's next location, based on previous object measurements. After developing the predictive filter, the third section discusses three methods of integrating the diffusion pyramid with the predictive filter. The appendix concludes by presenting results quantifying solution quality improvements attained by incorporating the anisotropic diffusion pyramid into a target tracking system.

Background

## Kalman Filtering

Taking advantage of the knowledge of an object's previous location, the Kalman filter can be utilized for target prediction [20]. This filter is a combination of filter and predictor, where the predictor estimates the location of the target at time $t$ given $t$-1 observations and the filter improves this estimate by accounting for measurement uncertainty and random drift. For an image tracking application, the Kalman filter is divided into two independent and identical systems which correspond to the horizontal and vertical directions of the current frame. Since these systems are identical, and

assumed independent, the following discussion will detail the implementation of the Kalman filter for the prediction of row location $i$.

Given a sufficiently high acquisition rate, the constant velocity, or *alpha-beta*, model of the Kalman filter may be used and velocity changes accounted for as random drift. With this assumption, the prediction of the target's next position is given by

$$\hat{i}_{t+1|t} = \hat{i}_{t|t} + \delta T \; \hat{v}^i_{t+1|t} \qquad (C.1)$$

where $\hat{i}_{t+1|t}$ is the prediction for time $t+1$ at time $t$, $\hat{i}_{t|t}$ is the filtered estimate of the observed location, $\hat{v}^i_{t+1|t}$ is the predicted velocity, and $\delta T$ is the time difference between the two observations. The estimate $\hat{i}_{t|t}$ provides a filtered representation of the observed location, accounting for measurement uncertainty, and is given by

$$\hat{i}_{t|t} = \hat{i}_{t|t-1} + \alpha_t \, (i^0_t - \hat{i}_{t|t-1}), \qquad (C.2)$$

where $\hat{i}_{t|t-1}$ is the predicted location at time $t-1$, $i^0_t$ is the actual observed location, and $\alpha_t$ is the filter gain. This gain determines the balance between the previous track history and new observation. If $\alpha_t$ is near one, then the observations are very reliable and the track history is ignored. In the case of strong measurement noise, $\alpha_t$ is set near zero.

The target velocity $v^i_t$ is modeled by

$$v^i_{t+1} = v^i_t + u^i_t, \qquad (C.3)$$

where $u^i_t$ is the velocity drift, or acceleration, and the prediction for the velocity component at time $t$ is

$$\hat{v}^i_{t+1|t} = \hat{v}^i_{t|t-1} + \beta_t \, (i^0_t - \hat{i}_{t|t-1}), \qquad (C.4)$$

where $\beta_t$ controls the effect of the new observation on the predicted velocity. If $\beta_t$ is near zero, the observations are deemed unreliable and the velocity thought constant.

## Computing the Kalman gains

One of the properties of the constant velocity model is that, assuming stationarity, the gains may be computed before the tracker is implemented. Since these gains converge quickly to constants, only a few computations are necessary, and computational efficiency is further improved. In precomputing the gains of the predictive filter, both gains are dependent on the variances of the noise processes and the state vector error covariance matrix.

Letting the state vector $X_t$ be defined as $X_t = \begin{bmatrix} i_t \\ v_t^i \end{bmatrix}$, containing the actual position and velocity of the target, the state vector of the predictor can be expressed as

$\hat{X}_{t|t-1} = \begin{bmatrix} \hat{i}_{t|t-1} \\ \hat{v}_{t|t-1}^t \end{bmatrix}$, and the state vector for the filter $\hat{X}_{t|t}$ constructed similarly. The error in

the predicted state vector is $X_t - \hat{X}_{t|t-1}$, and the error for the filter state vector is $X_t - \hat{X}_{t|t}$. Since these errors are stochastic vectors, they have covariance matrices, where the predicted state vector covariance matrix is

$$P_{t|t-1} = E\left[ \left( X_t - \hat{X}_{t|t-1} \right)\left( X_t - \hat{X}_{t|t-1} \right)^{\mathrm{T}} \right],\tag{C.5}$$

and the filtered state vector error covariance matrix is

$$P_{t|t} = E\left[ \left( X_t - \hat{X}_{t|t} \right)\left( X_t - \hat{X}_{t|t} \right)^{\mathrm{T}} \right].\tag{C.6}$$

In order to provide the minimum mean squared error prediction, $\alpha_t$ and $\beta_t$ are chosen to minimize $P_{t|t}$.

Assuming normality for both noise processes, measurement noise variance $\sigma_n^2$ and velocity drift noise variance $\sigma_u^2$, the solution that minimizes $P_{t|t}$ is [18]

$$\alpha_t = \frac{P_{t|t-1}^{11}}{P_{t|t-1}^{11} + \sigma_n^2} \tag{C.7}$$

and

$$\beta_t = \frac{P_{t|t-1}^{21} \delta T}{P_{t|t-1}^{11} + \sigma_n^2}. \tag{C.8}$$

For the constant velocity alpha-beta model, $P_{t|t-1}$ in (C.7) and (C.8) can be computed recursively as follows:

$$P_{t+1|t}^{11} = P_{t|t-1}^{11} + 2P_{t|t-1}^{12} + P_{t|t-1}^{22} - \frac{\left(P_{t|t-1}^{11} + P_{t|t-1}^{12}\right)^2}{P_{t|t-1}^{11} + \sigma_n^2}, \tag{C.9}$$

$$P_{t+1|t}^{12} = P_{t|t-1}^{12} + P_{t|t-1}^{22} - P_{t|t-1}^{12}\left(\frac{P_{t|t-1}^{11} + P_{t|t-1}^{12}}{P_{t|t-1}^{11} + \sigma_n^2}\right), \tag{C.10}$$

$$P_{t+1|t}^{21} = P_{t+1|t}^{12}, \tag{C.11}$$

and

$$P_{t+1|t}^{22} = P_{t|t-1}^{22} + \sigma_u^2 - \frac{\left(P_{t|t-1}^{12}\right)^2}{P_{t|t-1}^{11} + \sigma_n^2}. \tag{C.12}$$

Initializing the Kalman filter

To realize the Kalman filter, the initial conditions for $P_{t|t-1}$ (= $P_{1|0}$) are needed.

For a description of these conditions two additional parameters: $\sigma_i^2$, the variance in the initial position and $\sigma_{v_i}^2$, the variance of the initial velocity are defined. Assuming that the initial position is a uniformly distributed random variable, the calculation of $\sigma_i^2$ is straightforward. The computation of $\sigma_{v_i}^2$ can be derived using *a priori* knowledge of the target velocity range and distribution.

The filtered state vector error covariance at time 0 is given by

$$P_{0|0} = E\left[(X_t)(X_t)^{\mathrm{T}}\right] = \begin{bmatrix} \sigma_i^2 & 0 \\ 0 & \sigma_{v_i}^2 \end{bmatrix}. \tag{C.13}$$

$P_{1|0}$ can be computed using

$$P_{1|0} = A_0 P_{0|0} A_0^{\mathrm{T}} + Q_0 \tag{C.14}$$

where $A_0$ is the state transition matrix, and

$$A_0 = \begin{bmatrix} 1 & \delta T \\ 0 & 1 \end{bmatrix}. \tag{C.15}$$

$Q_0$ is the covariance of the system noise, and is given by

$$Q_0 = \begin{bmatrix} \sigma_n^2 & 0 \\ 0 & \sigma_u^2 \end{bmatrix}. \tag{C.16}$$

To initiate the tracker, $\hat{i}_{0|0}$ is set to the first acquired position. The first velocity estimate is indeterminate and can be set to any possible velocity.

Integrating the pyramid and Kalman filter

The Kalman filter is implemented using the following method. For initialization,

$P_{1|0}$ is computed and used to precompute the gains, $\alpha_t$ and $\beta_t$, for each $t$. Next, the target position is acquired and the filtered position, $\hat{i}_{t|t}$, is computed. This filtered location is then used to obtain the predicted velocity, $\hat{v}^i_{t+1|t}$, which allows the computation of the predicted position, $\hat{i}_{t+1|t}$. Once the next target position is predicted, it can be used to restrain the object search within the next scene.

The simplest view of a multi-resolution tracking system can be conceptualized as two independent systems, the image pyramid and tracking predictor. Within this context, the pyramid is generated with no knowledge of the predictive information from the tracker. The system is initialized by a multi-resolution search which locates the object of interest and passes its initial location to the tracker. Besides the initial acquisition of the target, the image pyramid is not constructed and the search performed on the highest resolution image. Given a reliable tracking prediction, this system would provide significant performance enhancement, since the cost of pyramid construction is restricted to the initial target search. Unfortunately, the robustness of the system is dependent solely on the accuracy of the tracking model and does not take advantage of the multi-scale recognition techniques. Additional inefficiencies are realized in practical implementations as the original signal must be preprocessed to remove noise. Given a filter which satisfies the prefiltering requirements of a decimation operation, the search becomes inefficient as redundant information is processed.

If the two systems are allowed to interact then further gains are possible. Since the prediction mechanism has its own uncertainty attribute, the root level of the search should be redefined to account for the accuracy of the prediction. In the event that the

predictive uncertainty is low, the root level would be lowered and the search constrained to finer scene representations in the pyramid. Alternatively, if the prediction uncertainty is high, then the root level would be defined as it was originally. In this case, the initial search would begin at the original root level, but search only a small region. Both cases allow the multi-resolution search additional refinement, increasing computational efficiency and system robustness.

Besides constraining the search, the predictive information could also constrain the pyramid and further increase computational efficiency. Given a prediction with known uncertainty, the image pyramid will only be searched within a region defined completely by the predicted position and the root level of the search. This does not require construction of a full image pyramid, but rather one centered at the predicted information and of size determined by the object size and root level. This optimization provides further computational efficiency, assuming the pyramid generating mechanism can be utilized for other purposes, without sacrificing solution quality. The actual performance of these different integration methods are presented in the next chapter.

## Results

To provide a quantitative evaluation of system performance, the anisotropic diffusion pyramid was applied to a sequence of 25 infrared images of a jet airplane in flight. Here, results are presented which compare the multi-scale techniques to single resolution methods for both target recognition and tracking applications.

The pyramids used for this evaluation are constructed with a 1 of 2 uniform sampling scheme, a gradient threshold of 15, and a $\Delta t$ of ¼. The original infrared images. and the base of the pyramid, have a resolution of 256x256 pixels with each pixel able to represent 256 intensity levels. Figure C.1 shows the pyramid constructed for the first frame of the sequence. To implement a multi-resolution search, the root level of the object must be identified. In the entire sequence, the largest element of the aircraft is its fuselage, and using the method discussed in Chapter 6, the root level of the sequence is defined as the third resolution representation above the original image.
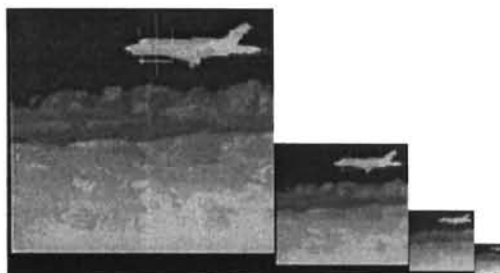


**Figure C.1** The first four levels of the anisotropic diffusion pyramid for frame #1 of the original sequence.

Applying multi-resolution techniques to the object recognition problem and using the third level of the pyramid as the root level, the object recognition tasks were performed using a binary edge based template matching routine. For a single resolution match, the algorithm required approximately 172 seconds per frame on a Sun Ultra 1/170. The multi-resolution technique required approximately 8 seconds per frame. The results show a system performance improvement of 21 times tradition single-resolution methods.

Besides providing computation efficiency, multi-resolution techniques also increase system robustness. Using the same binary template matching routine, pixel localization errors were computed for both single and multi-resolution trials. These results are summarized in Table C.1, presented at the end of this appendix, where the localization error is expressed as the Euclidean distance between the observed point and

ground truth. For the first 14 frames, the algorithms produce similar measurements. In the final 11 images, the multi-scale search is able to locate the target, while the single-resolution method is not.

To display robustness, the simulations were performed on the same set of images, but corrupted by Gaussian distributed noise. Figure C.2 shows the pyramid constructed for the first frame of the noisy sequence. The mean-square signal to noise ratio of the test images was approximately 15.7. As can be seen from Table C.1, the pixel localization error of the multi-scale technique increases, but the algorithm is still capable of estimating the object location.
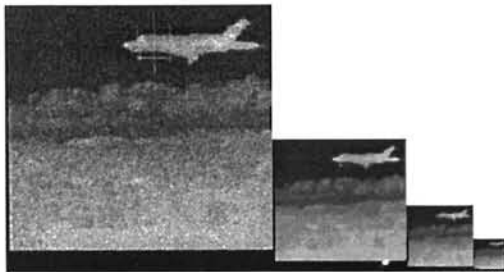


**Figure C.2** The first four levels of the anisotropic diffusion pyramid for frame #1 of the corrupted sequence.

The single resolution method is unable to locate the object during any frame of the sequence. The ability to find an object in high clutter allows the multi-scale object recognition system to provide a smaller pixel localization error with mean error of 3.69 pixels compared to 195.29 pixels of the single resolution system.

After displaying the efficiency and robustness of the multi-scale techniques to object recognition, results were then computed for a variety of tracking systems. Previously, three techniques were suggested for integrating a multi-scale structure and a predictive tracking filter: a multi-scale search followed by single resolution inspections, a multi-scale search followed by multi-scale examinations, and a multi-scale search followed by the inspection of a constrained pyramid. The results in Figure C.3 compare the localization errors for the single resolution approach to the three multi-resolution

systems using both the original and corrupted sequences. Figure C.4 compares the

computational requirements of the single resolution and multi-resolution approaches.
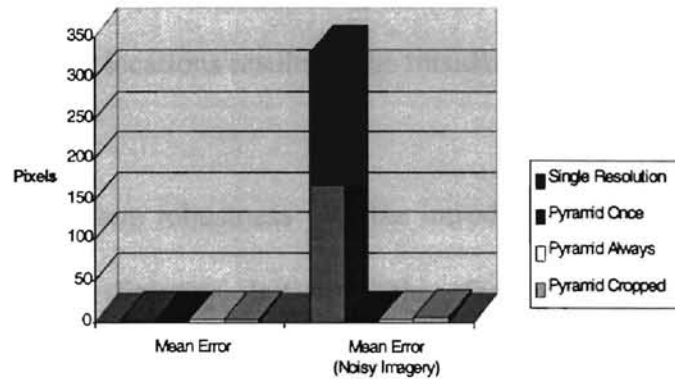


**Figure C.3** Comparison of the multi-scale tracking methods. Object localization errors are the Euclidean distance between the observation and ground truth.
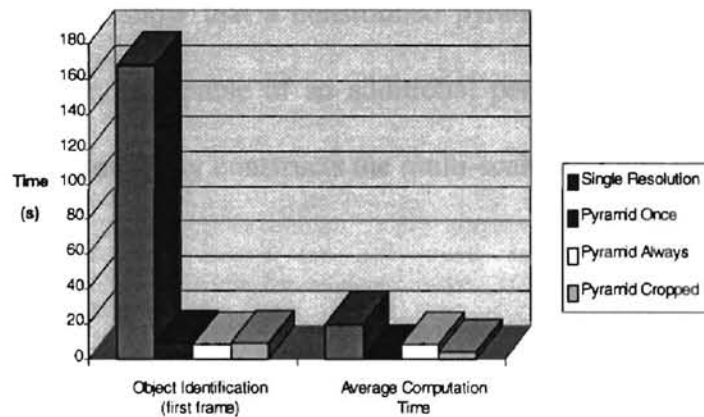


**Figure C.4** Computational requirements of the multi-scale tracking methods.

The results of the tracking simulations initially show the single resolution method

as the most accurate solution to the tracking scenario. To achieve this performance, the

algorithm was constrained to inspect only nine possible object locations within a frame.

The prediction of these locations is the responsibility of the tracking filter. Since the

original test data consists of well behaved constant velocity motion, these estimates are

accurate, and the system performs reliably and efficiently. Comparatively, the pyramidal algorithm is capable of searching 729 possible object locations while increasing the localization error by only 3%. Allowing the single resolution tracking system the same number of possible target locations results in the misidentification of the target within the first three images.

Increasing the system robustness with the introduction of an image pyramid, also increases computational efficiency. Initial searches, for this data set, were computed 20 times faster with a multi-scale structure. Subsequent inspections on equivalent search areas were performed over two times faster than single resolution techniques. This performance is further increased with the localization of pyramid construction to a smaller region. Results show that a constrained pyramid, simulating large uncertainty from the tracking filter, is capable of an additional performance increase of two times over a system which completely constructs the multi-scale structure for each frame.

## Localization Errors

| Frame Number | Multi-Scale | | Single Resolution | |
|---|---|---|---|---|
| | Original | Noise | Original | Noise |
| 1 | 4.47 | 9.06 | 4.47 | 207.76 |
| 2 | 3.35 | 5.22 | 3.35 | 206.71 |
| 3 | 3.16 | 2.24 | 3.16 | 219.59 |
| 4 | 2.83 | 2.24 | 2.83 | 211.15 |
| 5 | 1.41 | 1.00 | 1.41 | 208.81 |
| 6 | 1.41 | 2.24 | 1.41 | 200.35 |
| 7 | 1.12 | 4.72 | 1.12 | 199.20 |
| 8 | 1.00 | 1.41 | 1.00 | 209.85 |
| 9 | 0.50 | 1.80 | 0.50 | 208.63 |
| 10 | 0.50 | 1.12 | 0.50 | 206.59 |
| 11 | 1.12 | 3.50 | 1.12 | 194.16 |
| 12 | 1.12 | 2.06 | 1.12 | 192.55 |
| 13 | 1.00 | 4.24 | 1.00 | 195.73 |
| 14 | 1.12 | 3.04 | 1.12 | 192.49 |
| 15 | 2.06 | 2.69 | 159.00 | 189.18 |
| 16 | 1.12 | 5.32 | 1.12 | 194.31 |
| 17 | 1.12 | 4.27 | 155.98 | 186.60 |
| 18 | 1.12 | 3.64 | 1.12 | 185.50 |
| 19 | 3.35 | 6.73 | 162.65 | 185.58 |
| 20 | 2.50 | 4.61 | 175.02 | 184.29 |
| 21 | 1.12 | 5.41 | 159.89 | 182.77 |
| 22 | 2.50 | 5.32 | 166.90 | 180.13 |
| 23 | 3.35 | 2.69 | 154.84 | 180.57 |
| 24 | 3.61 | 1.41 | 149.05 | 180.23 |
| 25 | 3.61 | 6.40 | 99.82 | 179.61 |
| **Mean** | **1.98** | **3.69** | **56.38** | **195.29** |

**Table C.1** Pixel localization errors for multi-scale and single resolution object identification algorithms. Errors are expressed as the Euclidean distance between observed location and ground truth.

147

VITA

C. Andrew Segall

Candidate for the Degree of

Master of Science

Thesis: CONSTRUCTION OF THE SCALE AWARE ANISOTROPIC DIFFUSION PYRAMID WITH APPLICATION TO MULTI-SCALE TRACKING

Major Field: Electrical Engineering

Biographical:

Education: Received Bachelor of Science degree in Electrical Engineering from Oklahoma State University, Stillwater, Oklahoma in December 1995. Completed requirements for the Master of Science degree at Oklahoma State University in December, 1997.

Professional Experience: Electronic Media Computer Specialist; Electronic Media Development, Oklahoma Department of Vocational and Technical Education, 1990 to 1996; Graduate Research Assistant, Oklahoma Imaging Laboratory, School of Electrical and Computer Engineering, Oklahoma State University, 1996 to present.