WARWICK
THE UNIVERSITY OF WARWICK

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:
http://wrap.warwick.ac.uk/109834

Copyright and reuse:
This thesis is made available online and is protected by original copyright.
Please scroll down to view the document itself.
Please refer to the repository record for this item for information to help you to cite it.
Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications

THE BRITISH LIBRARY DOCUMENT SUPPLY CENTRE

TITLE MULTIRESOLUTION IMAGE MODELLING AND ESTIMATION

AUTHOR Simon Clippingdale B.Sc.

INSTITUTION and DATE 1988 The University of Warwick

Attention is drawn to the fact that the copyright of this thesis rests with its author.

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no information derived from it may be published without the author's prior written consent.

THE BRITISH LIBRARY
DOCUMENT SUPPLY CENTRE
Boston Spa, Wetherby
West Yorkshire
United Kingdom

REDUCTION X 20

CAMERA 8

MULTIRESOLUTION IMAGE MODELLING AND ESTIMATION

Simon Clippingdale B.Sc.

A thesis submitted to
The University of Warwick
for the degree of
Doctor of Philosophy

September 1988

Multiresolution Image Modelling and Estimation

Simon Clippingdale B.Sc.

A thesis submitted to
The University of Warwick
for the degree of
Doctor of Philosophy

September 1988

Summary

Multiresolution representations make explicit the notion of scale in images, and facilitate the combination of information from different scales. To date, however, image modelling and estimation schemes have not exploited such representations and tend rather to be derived from two-dimensional extensions of traditional one-dimensional signal processing techniques. In the causal case, autoregressive (AR) and ARMA models lead to minimum mean square error (MMSE) estimators which are two-dimensional variants of the well-established Kalman filter. Noncausal approaches tend to be transform-based and the MMSE estimator is the two-dimensional Wiener filter. However, images contain profound nonstationarities such as edges, which are beyond the descriptive capacity of such signal models, and defects such as blurring (and streaking in the causal case) are apparent in the results obtained by the associated estimators.

This thesis introduces a new multiresolution image model, defined on the quadtree data structure. The model is a one-dimensional, first-order gaussian martingale process causal in the scale dimension. The generated image, however, is noncausal and exhibits correlations at all scales unlike those generated by traditional models. The model is capable of nonstationary behaviour in all three dimensions (two position and one scale) and behaves isomorphically but independently at each scale, in keeping with the notion of scale invariance in natural images.

The optimal (MMSE) estimator is derived for the case of corruption by additive white gaussian noise (AWGN). The estimator is a one-dimensional, first-order linear recursive filter with a computational burden far lower than that of traditional estimators. However, the simple quadtree data structure leads to aliasing and 'block' artifacts in the estimated images. This could be overcome by spatial filtering, but a faster method is introduced which requires no additional multiplications but involves the insertion of some extra nodes into the quadtree. Nonstationarity is introduced by a fast, scale-invariant activity detector defined on the quadtree. Activity at all scales is combined in order to achieve noise rejection. The estimator is modified at each scale and position by the detector output such that less smoothing is applied near edges and more in smooth regions. Results demonstrate performance superior to that of existing methods, and at drastically lower computational cost. The estimation scheme is further extended to include anisotropic processing, which has produced good results in image restoration. An orientation estimator controls anisotropic filtering, the output of which is made available to the image estimator.

Key Words

Multiresolution, Image Modelling, Image Estimation, Quadtree

Acknowledgments

This work was conducted for the first year at the Department of Electronic Engineering, University of Aston and subsequently at the Department of Computer Science, University of Warwick. I should like to thank the staff and research workers of both Departments, and particularly the Image and Signal Processing Group at Warwick (Andy Calway, Roddy McColl, Ed Pearson, Martin Todd) for their interest and encouragement.

Thanks also to Professor Graham Nudd and the VLSI Signal Processing Group at Warwick for their patience during the final stages.

The design of the quadrature filtering scheme for orientation extraction is based largely on the work of Dr. Hans Knutsson.

The greatest debt I owe to my supervisor Dr. Roland Wilson, without whose ideas, enthusiasm and profound expertise in the subject this work would not have been possible.

CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introductory Remarks

Information-processing systems generally represent 'real-world' events or processes by the use of signals. A signal in this context is a function of a number of indices, the number being equal to the dimension of the signal space.

The transition from the actual, real event to its signal representation (for example the capture of a light distribution and its conversion to an electrical representation by a television camera, or the transduction by sensors of temperature or pressure into electrical signals) necessarily involves some input of energy from the environment. At this stage, and in subsequent processing, the signal is liable to distortions which render it a poorer representation of the original process.

Spurious energy in the environment, sensor noise and nonlinearities, and other distortions generated internally at later stages in processing combine to corrupt the signal. Physical, financial and technical constraints limit the degree to which such degradations may be avoided, and typically the information-processing system is bound to work with corrupted signals.

As an example, consider a satellite-mounted system which captures images and transmits them to a distant receiver. Distortions are introduced by thermal (Johnson) noise and nonlinearities in the sensor and transmitter, and by disturbances in the environment exciting the receiver along with the transmitted signal. This latter source of degradation could be tackled by increasing the transmitter power, but this has adverse implications for the

life of the satellite batteries. A similar problem obtains in medical X-ray applications, where a better image would often result from increasing the power of the X-ray source, but the danger to the patient prohibits such a solution. In both examples, the subsequent analyser of the images is compelled to accept corrupted data.

It may, however, be possible artificially to remove some of the degradation from the signal if the 'true' signal and the corruption effects can somehow be identified and separated. Clearly this requires a degree of prior knowledge about their respective properties or structures and furthermore that these structures be different. The design of systems for achieving such a separation falls within the field of signal restoration.

The goal of signal restoration is to obtain the best possible estimate of the true signal from the available data. The definition of 'best' here is crucial to the development of a restoration scheme and is by no means unique. Different definitions are appropriate in different circumstances, and the choice has far-reaching implications for the tractability, computational complexity and ultimate utility of a restoration scheme.

1.2 The Estimation Problem

The restoration task involves the estimation of the 'true' signal (henceforth referred to simply as the signal) from a volume of available observed data[7][25][99]. As noted above, it is necessary that the signal and the corruption (which will mean the difference between the signal and the data) be in some sense distinguishable.

Assuming that the signal carries useful information, it must be to some extent unpredictable and hence its variation in each of its dimensions cannot be exactly known. The same will be true for the random components of the corrupting process.

Accordingly, one is compelled to resort to more general statistical descriptions of the signal and of the various degradations. These descriptions are referred to as *models*. Since the estimation scheme will be designed on the basis of such models, their realism — that is to say, how faithfully they describe their respective actual processes — will affect strongly the ability of the scheme to achieve its goals.

Returning now to the question of what constitutes the 'best' estimate of the signal, it is natural to define as a starting point the *estimation error*, which is itself another signal, given as the difference between the 'true' and estimated signals. Clearly, were the estimation error to be identically zero, then the estimator would be unimprovable, and so the error signal is a meaningful quantity. It is then possible to define a *cost function* which expresses the penalty or undesirability associated with a given value of the estimation error.

It is here that various estimation strategies diverge. For example, the cost function may be simply the absolute value or the square of the error; it may be the mean of either of these quantities; it may be the maximum of either; it may depend also on the signal value such that the same error is more or less significant at different signal values, or it may depend on any number of functions of the error and of the signal. The *raison d'être* of the cost function, however, is that it is the minimisation of this function which yields the optimal estimator for the chosen signal and corruption models and the chosen cost function.

The choice of the cost function is determined by what is known as the *observer model*. This is a description of the significance of the estimation error signal to the 'end-user' or *observer* of the restored signal. The observer model might, for instance, attach a large significance to individual isolated large errors if the observer were a fast-responding air-

craft flight control system driven by restored trajectory estimates. In the same case, an overestimate of altitude might be far more dangerous and undesirable than an underestimate, and either might be more significant at very low altitude. If the signal were an image such as a medical X-ray and the observer an automatic 'tumour detector', the failure to detect a tumour might be weighted more heavily than a false alarm. If the signal were a natural image and the observer a human visual system, the observer model might attempt to include the known sensitivity of the visual system[39][91] to oriented features such as lines and edges, and its relative insensitivity to noise in these areas. An even more sophisticated variant might include the particular perceptual tuning which man possesses for human faces and especially eyes.

As the above discussion hints, there is a danger that ever more sophisticated and realistic observer models can lead to complex cost functions which may render the associated optimal estimator computationally unfeasible or even totally intractable. This is especially so for observers such as the human visual system which are very complicated in their sensitivity to errors.

Primarily as a result of its simplicity and tractability, by far the most commonly used cost function for estimation schemes has been *mean squared error* (MSE)[7][25][93][99], where the cost function to be minimised is the expectation over the signal probability space of a quadratic in the estimation error. Thus the optimality criterion (which is simply a statement of the objective of minimising the cost function) in this case is *minimum mean squared error* (MMSE). This criterion is also known occasionally as the *r.m.s.* (root mean square) error criterion, since minimisation of the square root of the MSE is equivalent, or as the *Wiener* criterion[84] after one of the pioneering contributors to the field[143]. The term *least-squares estimation* is also used, being somewhat more easily enunciated than 'minimum mean squared error estimation'.

For the class of quadratic cost functions and a number of others, it may be shown[7][99][110] that the optimal estimator of the signal from the available data is simply the conditional mean (expectation over the governing probability space) of the signal, conditioned upon the data.

The conditional mean is in general a nonlinear function of the data, but in the important case where the data obey a jointly gaussian probability density function (pdf), the conditional mean may be shown[110] to reduce to a linear form. The associated optimal MMSE estimator is then a linear combination of the data, and the procedure is known as *linear minimum mean squared error* (LMMSE) estimation. Note that it is always possible to construct a linear estimate, but only for jointly gaussian processes will this estimate be optimal in the MMSE sense.

1.3 Image Models

Over the last 25 years or so, a number of attempts have been made to model images with a view to applications in restoration, enhancement and above all in coding. The very considerable redundancy (correlation) which exists in natural images[63] and to an even greater extent in sequences of images such as television pictures[105] suggested that models which captured this property implicitly or explicitly might yield huge reductions in the volume of information which was required to be transmitted. This implies a commensurate reduction in the requisite bandwidth and so permits more television stations per unit of electromagnetic spectrum.

Differential pulse code modulation (DPCM) coding applications[40][42][63] have usually taken the image in raster-scanned format and treated it as a one-dimensional signal with stationary statistics and high correlation between adjacent pixels. This model is implicit

rather than explicit in the heuristic design of schemes such as delta modulation (DM)[41]. The model breaks down at sharp luminance discontinuities such as edges and the DM coder may saturate, but these effects were tolerated in view of the simplicity of the coder structure. Vertical correlations in the image were ignored.

The raster scan imposes *causality* on the image model. Causality implies a (time) direction which defines 'past', 'present' and 'future' regions of an image. At the present pixel, one may make use only of data from the present and past, since it is assumed that data from the future is unavailable.

A class of causal models, borrowed from the physics of Brownian motion, is the causal autoregressive (AR) or Markov process[35][86][110][138]. The present pixel is modelled as a one-dimensional linear combination of preceding pixels (which represents the minimum-variance predictor[86] of the current pixel) and a white noise term; the combination may be expressed as a finite difference equation known as the *innovations representation* [65][68][69] of the signal. This model leads to predictive coders with a white prediction error of minimum variance (i.e. minimum mean squared error (MMSE)), which is theoretically optimal for coding purposes[63]. The innovations representation of the model leads directly to causal recursive estimation strategies[67][69] which are linear shift-invariant (LSI) under assumptions of gaussian behaviour (hence linear) and stationarity (hence shift-invariant) in the statistics of the causal prediction error. The number of preceding pixels which are included in the difference equation defines the *order* of the model. The higher the order, the greater will be the amount of computation required at each pixel. The model order reflects the distance over which correlations are presumed to exist in the signal.

A generalisation of AR signal models is the autoregressive moving-average (ARMA)

process[18][110][133][138] which extends the AR process to the case where the driving term is a moving-average, coloured noise. Zeroes may appear in the spectral density function (SDF) of ARMA processes[110] in addition to the poles contributed by the AR component.

The above one-dimensional causal models may be extended to the noncausal case, where data from the 'future', in the sense described above, is available. This obviously has implications for the 'real-time' applicability of such schemes, since at least a delay would need to be introduced.

Noncausal, one-dimensional AR models[62][63] express the current pixel as a predictive term which is given by a difference equation involving 'past' and 'future' pixels, combined with a prediction error term as in the causal case. Since more data is available to the predictor, the prediction error variance (i.e. the MSE) is smaller than for the causal model, but now the error term is non-white[63] and no innovations representation exists. The consequence of this is that there is not a recursive implementation for the predictor/estimator. This computational disadvantage often far outweighs the lower MSE achieved by noncausal prediction models.

In applications where the processing system has access to at least a few lines in advance of the current pixel and perhaps to the entire image, the restriction to one-dimensional models is fatuous since the signal is two-dimensional and is available in that form. Causality is a particularly incongruous notion in the case of two-dimensional signals such as images, which have no clearly-defined dominant axis or direction. However, many established and effective signal processing techniques were developed for one-dimensional signals, and there has been a desire to extend their use to image data in expectation of similar benefits. This accounts for the attempts which have been made to

extend the above models — including one-sided and causal structures — to two dimensions.

The AR models in two dimensions naturally use a two-dimensional prediction window. The geometry of this window defines the model type.

If only 'past' pixels (in the sense of the causality imposed by the raster scan) are included, the prediction window is termed *non-symmetric half-plane* (NSHP) [27][29][36][157]. The predictor may use any pixels from lines above the current line, and pixels on the current line prior to the present pixel. Driven by a noise process, the NSHP model generates a class of Markov random field (MRF), although Woods[157] remarks that not every MRF may be decomposed by spectral factorisation into a finite-order NSHP model. The NSHP model leads to recursive estimators[28][60][97][156] which are two-dimensional extensions of the well-established Kalman filter[69][70][71] for one-dimensional signals.

If the prediction window includes pixels from lines above the current line, and the current line itself with the exception of the present pixel, the model is termed *semicausal* since it has causal structure in one Cartesian direction (the vertical for conventional raster scan) and noncausal in the other. Jain[61][64] has developed difference equations and filtering schemes for models of this type.

The 'ultimate' two-dimensional prediction model is the full-plane, noncausal variety in which the prediction window potentially includes the entire image. Such a model leads to a general Markov random field characterisation[135][153][154]. The availability of the entire image admits the various transform domains[5][65][66] in its representation; the optimal solution for least-squares filtering given stationary image statistics may economi-

cally be expressed in the frequency (Fourier) domain as the classical Wiener filter[65][111][118].

Whilst as in the one-dimensional case such noncausal methods tend to achieve better results in terms of error measures (MSE or other), they do not admit of recursive implementations for the optimal predictor or estimator, and for computational reasons there remains much interest in such causal structures for image restoration [121][134][152][155] despite the apparent and intuitive unsuitability of causal image models.

1.4 Least-Squares Estimation of Images

As noted above, the least-squares or MMSE criterion provides a tractable solution to the linear estimation problem. This is so because differentiation and minimisation of the mean squared error with respect to the model coefficients yields a system of linear equations with an analytic and unique solution. The system is known as the *normal* or *Yule-Walker* equations[86][110], and requires for its solution a knowledge only of the second-order moments[110] (i.e. the correlation structure) of the signal and data joint probability density functions.

Thus an assumed or estimated correlation function for the signal and for the corrupting effects is sufficient to specify the optimal linear MMSE estimator, which will be identical to the (generally superior) nonlinear MMSE estimator if the variates are gaussian. The tractability of linear schemes by comparison with their nonlinear counterparts motivates the development of linear estimators even for non-gaussian processes; an assumption of gaussian behaviour may be implicit if a linear estimator or predictor is described as optimal rather than sub-optimal.

The application of least-squares estimation to images, by way of Wiener filtering, dates from the late 1960s and early 1970s with the work of Slepian[128], Helstrom[47], Pratt[118], Huang et al[52] and Hunt[57].

Nahi[102], Habibi[43], Nahi and Assefi[98] and Powell and Silverman[117] extended the recursive methods of Kalman filtering to two dimensions. Nahi and Habibi[101] introduced the first 'multiple-model' approach (see section 2.1) in an attempt to overcome the blurring of edges which is a consequence of linear shift-invariant (LSI) filtering. Nahi and Franco[100] developed the vector scanning model and its associated Kalman filter.

Since that era, a number of more sophisticated image modelling and estimation schemes have been introduced. Consideration of these will be deferred until Chapter 2, but it suffices to say that they may still be classified into essentially two approaches:

(i) causal models which yield causal and recursive estimators, and

(ii) noncausal models which yield noncausal and nonrecursive estimators.

1.5 The Nature of Images and of Vision

Natural images exhibit various general properties which are worthy of note, and which provide some indication of what might constitute desirable properties of image models. Such properties derive from a consideration of the properties of the material world which is represented in such images.

Firstly, the world, and hence the visual environment, is composed in general of objects; that is to say, it is a particulate rather than a diffuse system. That this should be so is a

consequence of the cohesiveness of matter.

A second property which follows from the first is that objects tend to have distinct and continuous boundaries. This observation may be encapsulated as the *convexity* of objects: matter does not tend to form agglomerates which are exotically-shaped with wildly irregular boundaries — rather it tends to prefer shapes which approximate the spherical.

From this property may be deduced another: the convexity property holds over a wide range of scales, and the agglomeration process is substantially similar in character over such a range. Thus develops the notion of *scale invariance*, which suggests that the first two properties may be observed to operate in the real world, and in its visual projection, over a wide range of scales. The concept of scale invariance in images implies that the structure of a natural image may be expected (in a general sense and over a reasonable range) to be rather similar whatever the size or scale of the portion of image under inspection. This notion will be taken up in more detail in Chapter 2.

The idea of the scale invariance of images may be invoked to support the contention that at a given scale, some objects will be relatively 'large' and therefore their boundaries will be approximately one-dimensional when viewed at this scale. This assertion leads naturally to the characterisation of edges in images as locally one-dimensional, with a well-defined position and *orientation* when viewed at an appropriate scale. Since the visual world may loosely be regarded as the sum of such region boundaries and their respective interiors (which are essentially isotropic by comparison with, and at the scale of, the edge) it may be postulated that such position and orientation information, when considered over a range of scales, might give a rather complete or at least functional description of the image.

The results of a number of physiological and psychophysical experiments[9][54][55][82] indicate that this information is indeed of some utility; it appears that the mammalian visual system operates at its lower levels in very much the manner suggested, with detectors tuned for particular orientations and scales of structure (i.e. edges) in the retinal image.

The work on artificial neural systems of Linsker[85] and others has also shown, rather surprisingly, that a simulated layered network of elements possessing only the barest functional resemblance to visual neurons will, when trained with a set of input patterns (even random noise), develop detectors with size and orientation tuning. The learning rule is simple and essentially maximises the variance or information at the output of a given element. This is not to imply that vision necessarily works in this way, but it does suggest that there is valuable information in such a representation.

1.6 Limitations of Least-Squares Image Estimation

As noted in section 1.3, many image models are defined on causal structures, and this restriction while leading to recursive least-squares estimation schemes is not of *prima facie* suitability in representing image data.

Furthermore, the assumption of stationary statistics while again simplifying the estimation strategy is clearly violated by the presence in images of edges. This mismatch accounts for the blurring of edges which is a consequence of shift-invariant processing.

The discussion of the choice of estimator cost function in section 1.2 indicates that the optimality criterion of the visual system is somewhat complicated in its sensitivity to errors. In particular, errors are known[105][129] to be less visible in regions of high

intensity or high intensity gradient. In sharp contrast, the least-squares (MMSE) criterion is non-contextual and weights equally any error of a given magnitude, irrespective of where in the image it occurs and of the behaviour of the image in the locality.

It is not surprising given the major differences between the two optimality criteria that estimation schemes which are optimal under the MMSE criterion perform rather badly when evaluated subjectively, i.e. by the visual system of a human observer.

The above discussion leads to something of a dilemma in the design philosophy of a restoration scheme. One may either persist with linear shift-invariant least-squares estimation in the knowledge that the resulting computational structure will be simple and tractable but will produce less than excellent visual results, or one may abandon least-squares estimation in favour of more complicated strategies or heuristic approaches which make no claim of optimality but are justified in an *ad hoc* manner by the subjective quality of their results.

The median filter[34][53][104] is one example of a nonoptimal but effective 'ad hockery' which suppresses large excursions in the data by assigning to its output the local median value of the data within its support. The filter is clearly nonlinear, but for modest window (support) dimensions is computationally manageable and tends by virtue of its nonlinearity to avoid the blurring of edges in images which is associated with linear filtering.

As section 1.2 suggests, it may be possible to obtain *optimal* estimation strategies which are better suited to the properties of both images and observers by utilising a more sophisticated error cost function. The problem here is the complexity or even intractability of the associated optimal estimator. It may however be possible to incorporate the added complexity into the signal model, and then to use a relatively simple error criterion

such as least-squares. The result should be a tractable but more appropriate estimator.

This proposition underlies the development of the 'signal-equivalent' modelling approach adopted by Abramatic and Silverman[2][3] and extended by Knutsson, Wilson and Granlund[75], which will be considered further in Chapter 2.

1.7 Principal Modifications to the Least-Squares Method

Bearing in mind the points raised in the preceding sections, it is possible to enumerate a minimal set of modifications to the conventional shift-invariant least-squares technique which might render it better suited to the problem of image restoration. Such modifications include the following:

(i) Noncausality — as has been pointed out above, the notion of causality is alien and inappropriate for the modelling of *image data*. (It is actually a model of a particular acquisition method, not of the data itself.) However, there is no question that the recursive filter implementations which follow from many causal models are desirable; one would nevertheless prefer a 'fast' filtering method based on a noncausal scheme if such could be developed. Noncausal approaches tend to be based on Wiener filtering and are computationally rather expensive, requiring at each pixel a number of multiply-accumulate operations equal to the size of the filter mask or equivalently forward and inverse two-dimensional discrete Fourier transforms.

(ii) Nonstationarity — linear shift-invariant (LSI) filtering takes no account of precisely those image details such as lines and edges which are known to be of considerable saliency in the cost function or optimality criterion of the visual system. Some degree of spatial adaptivity has long been recognised as vital if such features are adequately to be

preserved, and so a spatially variant strategy is a fundamental requirement. The computational cost of shift-variant processing must, however, be taken into account. The nonstationary and noncausal Wiener technique of Abramatic and Silverman[3] is an example where the computational burden is reduced by the deployment of the signal-equivalent model formulation.

(iii) Scale Invariance — as noted in section 1.5, scale invariance is a property of natural images, but it does not fit easily into the structure of many existing image modelling or estimation schemes. A strategy capable of incorporating this property would be desirable.

(iv) Computation — section (ii) mentions the generally computation-intensive nature of nonstationary schemes. This is compounded in the case of noncausal approaches by the lack of a recursive implementation. Unfortunately, two of the features listed above as desirable are noncausality and nonstationarity. Thus the development of a restoration strategy along the lines proposed should be guided also by the need to avoid the emergence of a computationally cumbersome implementation.

1.8 Thesis Outline

This work will treat the points raised thus far and attempt to provide workable solutions to some of the difficulties.

1.8.1 The Image Model

A new noncausal image model based on a multiresolution or pyramidal data structure will be introduced. The model is naturally scale-invariant in form and handles small- and

large-scale structure in identical ways. An alternative view is that at a given scale, objects constitute large-scale structure whilst their boundaries constitute small-scale structure. The model accounts for each type of data.

1.8.2 The Estimator

The least-squares optimal estimator associated with the model is developed for the case of corruption by additive white gaussian noise (AWGN). The estimator emerges in the form of a one-dimensional causal first-order recursive filter which has a very fast implementation, despite the fact that the image model itself is noncausal in the usual sense. The estimator is easily extended to the nonstationary case where the operator is modified according to local image activity. A fast, scale-invariant edge or 'activity' detector provides the contextual information.

1.8.3 Generalisation to Vector Data

The model and estimator are extended to the case where the input data is a vector-valued field with two indices, and is shown to be applicable to N-dimensional vector fields (i.e. fields with N indices). The scheme is applicable to vector restorations such as of multispectral data or parametric texture descriptors.

1.8.4 Further Modifications Motivated by Vision

The vector estimator is applied to the restoration of local orientation information computed from noisy images. The restored orientation data is then used to introduce local anisotropy into the estimation strategy.

1.9 Experimental and Display Conditions

'Real-world' images (original, noisy and restored images) are truncated to the eight bit integer values (0 to 255) used by the source material and the framestore. This means that for noisy images which contain values below zero or above 255 due to noise excursions, these values are displayed as 0 or 255 respectively. Note that this applies for display purposes only; all arithmetic was performed in single- or double-precision floating-point as appropriate.

This does, however, imply that the noisy images would appear somewhat worse were their full range to be displayed.

White gaussian noise was generated by the polar method[74] using the nominally uniformly-distributed output from the UNIX† random number generator *random*()[139]. The results were photographed with a Dunn Instruments Multicolor unit on Kodak TMAX 100 ASA monochrome film on the green channel with a nominal exposure time of 2.36 seconds.

All photographic figures are prefixed with 'P' (for example figure P23), and are situated at the end of the thesis. Other figures are indexed by chapter (for example figures 5.1 to 5.4 in chapter 5) and are located at the end of each chapter.

† UNIX is a trademark of AT&T Bell Laboratories.

CHAPTER 2

QUADTREE STRUCTURE AND NEW IMAGE MODEL

2.1 Motivation

As noted in Chapter 1, the minimum mean squared error (MMSE) criterion provides a tractable basis for the design of optimal estimation schemes. It is a goal of the present work that such MMSE optimality be retained, but in a form of estimator (and underlying image model) which is better suited to the problem of image restoration than the models discussed in Chapter 1.

A noncausal model and estimator are sought on the grounds that:

(i) Causality is not an appropriate attribute for image models if it is not mandated by 'real-time' considerations such as arise in on-line raster-scanned applications;

(ii) Noncausal estimators based on noncausal signal models achieve lower mean squared error than do corresponding causal models for a given class of signal. This results from the fact that the estimator is able to use 'future' as well as 'past' input values and so has more data on which to base the estimate.

Causal estimators, however, frequently give rise to recursive structures (such as the Kalman filter[69]) which are computationally more efficient than their noncausal counterparts. The causal models on which they are based often admit the so-called *innovations representation* [68] which is the key to the coercion of the resulting estimator into the recursive form.

Another desirable property of image models is nonstationarity. Natural images contain profound statistical nonstationarities such as lines and edges, and a realistic image model should account for this fact.

A stationary image model under stationary degradation (blur or noise) will give rise to a shift-invariant estimator which necessarily performs the same operations at every point in the image. This is clearly less than ideal in the context of natural images which tend to contain smooth regions with sharp edges, since a spatially-invariant restoration scheme which aims to smooth out noise is bound also to smooth salient linear features in the image.

Such features are known additionally to be of great importance to the visual system. Much evidence now exists[9][54][55][82] for the presence in the lower levels of the visual system of nerve cells (neurons) which function as linear feature (line and edge) detectors, implying that detection of such features constitutes an early and fundamental stage of visual perception. As a consequence, image modelling and estimation schemes which are based on stationary statistics are bound to produce visually unsatisfactory results.

The limitations of stationary modelling and estimation schemes suggest that a nonstationary approach is to be preferred. However, nonstationary schemes are apt to be computa-

tionally expensive since they generally involve the optimisation of the estimator at each point in the image.

This computational expense has led to the development of 'multiple-model' strategies which combine a fixed number of stationary models in a spatially-variant manner, in order to achieve some degree of spatial adaptivity without the need for the comprehensive recalculation at every image pixel of the model and its associated optimal estimator.

The multiple-model approach of Woods[152][157], for example, following the original work of Lebedev and Mirkin[83], uses five different stationary (Markov random field) causal models and switches between them on the basis of a controlling, higher-level model (a Markov chain) which governs the conditional probabilities of the transitions (see also [59]).

The computational and memory requirements are still rather demanding, since the method of [152] involves running a large number (25) of Kalman filters in parallel, each of which must carry forward its own state information; in addition, the selection of the currently operative model requires the calculation and sorting of the same number (25) of the *a posteriori* probabilities of each model, conditioned on the data, at each pixel.

Unfortunately, the parameters of both the higher- and lower-level models in [152] are estimated from uncorrupted versions of the source images, noisy versions of which are to be restored, and it is not clear how the system would perform given only the noisy data as would probably be the case in any practical application.

There is additionally a problem of system identification which restricts the ability of such causal multiple-model systems to switch quickly to the correct model, since the use of

the 'wrong' model has to continue for a finite period before it is possible reliably to detect that a change is required. The system of [152] exhibits this 'inertial' tendency for the currently-selected model to persist for some time after it has ceased adequately to represent the data.

Kashyap[72] has used another multiple-model approach, termed the 'multivariate random field', in which small segments (blocks) of an image are assumed homogeneous and each modelled by a stationary random field; the variation of the model parameters from one block to the next is described by a higher-level (vector) random field.

Whilst multiple-model approaches have achieved better results than the use of single models, the problems mentioned above restrict their effectiveness particularly in regions of abrupt and pronounced nonstationarity such as edges, and this is precisely where processing adaptivity is most required. The use of causal strategies is bound to lead to poor results in edge regions, even in multiple-model schemes.

A more 'continuously variable' modelling approach was introduced by Abramatic and Silverman[3], who defined a 'signal-equivalent' model which aims to incorporate the linear feature sensitivity of the visual system into a modified image representation. The image is modelled essentially as a combination of stationary lowpass and nonstationary highpass components, and the optimal (Wiener) filter for the lowpass component is combined in a spatially-variant manner with an identity operator such that less smoothing is applied to the noisy image in edge regions.

Knutsson, Wilson and Granlund[75] extended this method by recognising that an obvious and defining attribute of linear features (and again one which is important to the visual system) is orientation, and that an anisotropic term may profitably be incorporated into

the estimation operator. This leads to an estimation strategy which is both nonstationary and anisotropic, matching more closely both the properties of natural image features and the known mechanisms of vision.

However, the further one moves away from a purely stationary image estimation scheme toward 'full' nonstationarity (as opposed to the limited nonstationarity of multiple-model approaches) and even anisotropy, the greater in general is the computational burden. In the present work, it is desired that 'full' nonstationarity be available, but that the associated computational overheads be kept to the bare minimum.

Another property of natural images which might profitably be incorporated in a more realistic image model is scale invariance (see section 1.5). The structure of the natural world often looks similar over a number of different scales[150]. (Consider, for example, the observation that in an electron micrograph of a particle surface one sees 'hills', 'valleys' and so on). Edges in particular may exhibit some degree of fractal structure[87], and thus it would seem appropriate for the model to possess such properties.

Finally, the 'end-user' of processed images is often the (human) visual system, and its fidelity criterion is very different from minimum mean squared error[10][27][88][124][129][132]. As noted above and in Chapter 1, the visual system places much more emphasis on oriented features such as lines and edges.

Associated with this phenomenon are effects such as the visual masking effect[13][105] which describes the reduced sensitivity of the visual system to noise in the vicinity of such features. It has also been shown[75] that directionally-filtered noise which is oriented parallel to a nearby feature is far less disturbing visually than the same noise oriented in the perpendicular direction.

Hence the fidelity criterion of the visual system is at least nonstationary and anisotropic, and probably difficult to encapsulate in a mathematical form which both provides a complete description of its properties and admits a tractable form of optimal estimator.

In the light of the above, a choice is forced in the design process (see section 1.6). Either the MMSE criterion is abandoned in favour of heuristic methods (such as the median filter[34]) or, if it is desired that the basis of optimality be retained, the error criterion or equivalently the signal model must be modified to incorporate some or all of the desired characteristics.

In this work, it is desired that the MMSE criterion be retained as the basis for optimisation. What is sought is a noncausal, nonstationary image model giving rise to a noncausal, nonstationary but fast (i.e. computationally efficient) optimal estimator which is capable of incorporating modifications suggested by visual system properties.

2.2 Quadtree Structure and Applications

The quadtree[58][136] is the simplest of a class of pyramidal data structures which have found increasing application in image processing and computer vision work. Its introduction by Tanimoto and Pavlidis[136] in 1975 was the first attempt to provide a data structure matched to the 'multiresolution' representation of images, which has become recognised as valuable in a number of application contexts.

The concept of the pyramid arises because versions of an image at successively lower resolutions require fewer spatial samples for their representation and are thus 'smaller'. If the planes corresponding to these different resolutions are visualised as being stacked up so that the plane of highest resolution (the original image) is at the base, then the

structure resembles a square-based pyramid. It is common for each higher level to be of half the linear dimension (one quarter of the area or number of nodes) of its predecessor. Note, however, that the term 'quadtree' specifically implies that each node is linked only with four nodes on the level below and with one on the level above — a more general structure in which lateral and/or other vertical interactions occur is not a quadtree.

In image coding, the multiresolution approach offers the possibility of separating in some sense those components of an image which are of different scales. Not only does this provide a substantial decorrelation of the data, but it allows 'coarse' features to be encoded economically and transmitted before the finer detail. Such 'progressive transmission' techniques[73] build up what may be an adequate partial form of the image at the receiver with minimal information transfer.

Wilson[148] has used a quadtree method for predictive image coding, where the data at a node is formed by simple averaging over the four nodes to which it is linked on the level below. Quantisation and transmission commences with the top level (a single node) and proceeds in a recursive manner, with the difference between the value at a node and the quantised value at its parent being encoded. This scheme illustrates the general philosophy of pyramid coders and indeed of the present work; the data structure permits the use of simple predictive (DPCM) coding without constraining the coder to be causal in the image plane as is the case in conventional DPCM systems.

Burt and Adelson[11] introduced a multiresolution data structure in which each pyramid level is generated by gaussian filtering of the preceding level, giving a series of increasingly 'lowpass' images. The difference taken between appropriately interpolated versions of two contiguous levels constitutes a bandpass (Laplacian-filtered) image which may be encoded for transmission should the degree of detail represented in the given

band be required. The representation is invertible in that the original image may be recovered exactly (neglecting quantisation) from the data in all of the levels of the pyramid.

Adelson et al[4] have more recently discussed variants on the pyramid coder, including hexagonal and quincunx geometries and the use of quadrature mirror filter (QMF) kernels for the generation of the pyramid. Good reconstruction quality at a respectable data compression ratio was reported.

Cohen, Landy and Pavel[21] among others have applied quadtree coding to binary (cartoon) images. Here the node value is a binary attribute, and a node is labelled with the attribute if any of the nodes to which it is linked on the level below possesses it. The coding scheme is similar to that of Wilson mentioned above except for the binary nature of the data. Sparse cartoon images are particularly well suited to this type of coding, since for much of the image there will be no activity and in such regions the coding procedure may terminate economically at a high level of the quadtree.

Another major application area is the field of object/background or texture segmentation.

Burt, Hong and Rosenfeld[12] have developed a pyramidal scheme which achieves segmentation by forming links between nodes on adjacent levels, the link strength depending on the similarity in the node values. These values may simply be gray levels in the case of compact object segmentations, or alternatively they may be descriptors computed from a texture field which is to be segmented. The method is iterative, and the link strengths are adjusted until the given (predetermined) number of classes emerge. A multitude of similar strategies and extensions have been developed by Hong and Rosenfeld[49], Hong, Narayanan, Peleg, Rosenfeld and Silberberg[48], Pietikainen and Rosen-

feld[114], Hong and Shneier[50], and Hong, Shneier and Rosenfeld[51] among others. The latter work deals with edge extraction by the use of edge data in such a pyramid rather than region interiors. Chen and Pavlidis[17] have used a pyramid data structure for their 'split and merge' segmentation algorithm. Cohen and Cooper[20] use a modified quadtree for their hierarchical texture segmentation method, which is based on a Markovian texture model. Peleg at al[112] have studied the properties of textures in multiresolution representations from the viewpoint of random fractals[87], by considering the rate of increase of the area of a texture surface as the measurement resolution increases.

Spann and Wilson[130] have developed a quadtree-based segmentation method. This first involves smoothing by forming the quadtree of the image by averaging, and then classification at an upper level of the quadtree using a 'local centroid' clustering algorithm[147]. The classes thus obtained are propagated downward, with the positions of the interclass boundaries being estimated at each level such that at the image plane, the boundaries are only one or two pixels wide. This work addresses explicitly the problem of uncertainty[146][149], which limits the performance of many signal and image processing techniques in that it defines a bound on the degree to which global properties, such as statistical or frequency-domain parameters, and local spatial properties may simultaneously be effective. In segmentation, for example, the uncertainty principle dictates that certainties in class membership and in position are incompatible and that a tradeoff must exist. The same principle limits the degree to which edges may be localised in the presence of noise; this problem will be addressed in a subsequent chapter.

Multiresolution methods provide a means of handling explicitly the constraints imposed by the uncertainty principle; processing may be carried out over a range of scales and the information combined in a useful manner as the segmentation method described above illustrates.

Ranade and Shneier[122] have used a multiresolution (quadtree) adaptive smoothing technique. A quadtree is formed by averaging, and the smoothing process seeks to preserve edges in the image by proceeding downward in the quadtree until a node is reached below which there is little activity in the image, and assigning to all nodes below it the average value at the given node. Active regions such as edges are therefore assigned an average computed over a small region. The method is, however, somewhat *ad hoc* and makes no claim of optimality.

The properties of pyramid generating kernels have been considered by Meer, Bauer and Rosenfeld[92] and more recently by Watson[142], Wilson and Spann[150], and Wilson and Calway[145]. These workers note that simple spatial averaging will lead in the absence of explicit lowpass filtering to distortion caused by *alias* components[120] induced by the reduction in sampling rate as the pyramid is ascended. This is undoubtedly true of *any* kernel which exhibits less than total (and unrealisable) attenuation at spatial frequencies above half of the newly reduced sampling rate[142]. The introduction of lowpass ('anti-alias') filtering involves a considerable computational overhead, however, and it will be shown in Chapter 4 that this precaution may in certain cases be unnecessary.

2.2.1 Quadtree Structure

The quadtree is a pyramidal structure in which image data is represented on a number of different levels. A given level l consists of $(n_l \times n_l)$ nodes (l, i, j) where $i, j, \quad 0 \leq i, j \leq n_l - 1$ are the position coordinates of the node within the level.

Each node (l, i, j) on level l is linked to four nodes $(l+1, i', j')$ on level $l+1$, where $p' = 2p + q$, $q \in \{0,1\}$, $p \in \{i, j\}$.

Level $l+1$ is conceptually below level l. The node (l, i, j) is known as the 'father' of the four nodes $(l+1, i', j')$ (its 'children'). Level $l+1$ comprises $(2n_l \times 2n_l) = (n_{l+1} \times n_{l+1})$ nodes.

Level 0 contains just one node, known as the 'root' of the tree; this is the node (0,0,0) . The bottom level is level Y and is of the same dimension as the image data (in the present case $n_Y = 512$ and $Y = 9$).

The structure is depicted in figure 2.1.

A data value at a node is denoted as, for example, $s_{l,i,j}$ or $x_{l,i,j}$. The term 'for all i, j ' will imply $0 \le i, j \le n_l - 1$ if the level index l is clear from the context.

2.2.2 Ancestor and Descendant Sets

For a given node (l, i, j) in the quadtree, two sets of related nodes may be defined.

The *ancestor set* is defined as

$$A_{l,i,j} = \{(k,p,q) : 0 \le k \le l \ ,$$

$$2^{l-k}p \le i < 2^{l-k}(p+1) \ ,$$

$$2^{l-k}q \le j < 2^{l-k}(q+1)\} \ , \qquad (2.1)$$

and the *descendant set* as

$$D_{l,i,j} = \{(k,p,q) : l \le k \le Y \ ,$$

$$2^{k-l} i \le p < 2^{k-l}(i+1) \quad ,$$

$$2^{k-l} j \le q < 2^{k-l}(j+1) \;] \quad . \tag{2.2}$$

Note that $(l, i, j) \in A_{l,i,j}$ and $(l, i, j) \in D_{l,i,j}$. The ancestor and descendant sets are illustrated in figure 2.1.

The set of all nodes in the quadtree may be denoted by $D_{0,0,0}$.

The lowest common ancestor (LCA) of two nodes (l, i, j) and (k, p, q) is given by

$$LCA(l, i, j; k, p, q) = (r, m, n) \tag{2.3}$$

where

$$(r, m, n) \in (A_{l,i,j} \cap A_{k,p,q})$$

and

$$(r+1, s, t) \notin (A_{l,i,j} \cap A_{k,p,q})$$

for any s, t. The symbol $\cap$ denotes the intersection set.

2.3 A New Image Model

An image model, based on the quadtree structure, is defined as

$$s_{l,i,j} = \sum_{\substack{(k,p,q) \\ \in A_{l,i,j}}} \beta_{k,p,q} \, w_{k,p,q} \quad , \tag{2.4}$$

$$\beta_{k,p,q} \geq 0 \quad ,$$

where $w_{k,p,q}$ has a gaussian probability density function (pdf) and

$$E w_{k,p,q} w_{r,m,n} = \delta_{k,r}\delta_{p,m}\delta_{q,n} \tag{2.5}$$

where $\delta_{x,y}$ is the Kronecker delta function.

The model characterises an image as being generated by a random process defined on the quadtree. The process operates 'vertically' down the tree starting at the root node (0,0,0) and expresses the image as a sum of gaussian random variables of variance $\beta^2_{\ldots}$.

The model of (2.4) may be expressed in the recursive form

$$s_{l+1,i',j'} = s_{l,i,j} + \beta_{l+1,i',j'} w_{l+1,i',j'} \quad ,$$

$$(l+1, i', j') \in D_{l,i,j} \quad ,$$

$$0 \leq l \leq Y-1 \quad , \tag{2.6}$$

with initial condition

$$s_{0,0,0} = \beta_{0,0,0} w_{0,0,0} \quad . \tag{2.7}$$

This representation of the model expresses a child node in terms only of its father and a gaussian variable. The form of the model is illustrated in figure 2.2.

The $\beta_{k,p,q}$ are the parameters of the model and in general vary with p, q (i.e. with spatial position on level k). For the present, attention will be restricted to the spatially invariant form of the model, where $\beta_{k,p,q} = \beta_k$. The spatially variant model will be

considered in Chapter 4.

2.4 Properties of the Model

2.4.1 Noncausality

The model (2.6) is a martingale[110] causal in the level index l. However, as a consequence of its vertical nature, there is no preferred causal direction in the horizontal (spatial) plane and in the context of the image (as generated at level Y of the quadtree) the model is noncausal.

There is no temporal order on the image pixels such as that imposed by raster scanning, which gives rise to spatially causal models and causal estimators such as the Kalman filter[69].

The equation (2.6) may be considered to be a causal minimum variance representation (MVR) [86] or innovations representation [68] for $s_{\ldots\ldots}$ and so the model retains the simple causal structure provided by these representations while not restricting the image $s_{Y,\ldots}$ to be a causal signal.

2.4.2 Scale Invariance

Data $s_{l,i,j}$ at a node (l,i,j) projects on to the image plane (level Y) at all pixels $(Y,p,q) \in D_{l,i,j}$ in the descendant set of (l,i,j). A feature at (l,i,j) thus projects on to 4^{Y-l} nodes (Y,p,q) in the image (see figure 2.1).

Since the model of (2.6) is clearly isomorphic at all levels, it generates features in the image-plane projection $s_{Y_{p,q}}$ at all scales. Thus the model may be seen to be scale-invariant.

The values β_l at higher (smaller l) and lower (larger l) levels of the tree control the generation of larger and smaller image features respectively. This relates in an obvious way to the notion of scale invariance in natural images[150] (see section 1.5). Marr[91] comments, indeed, that

> 'The spatial organisation of a surface's reflectance function is often generated by a number of different processes, each operating at a different scale.'

The model is formulated in such a way as to incorporate not only the basis of scale invariance, but also this notion of 'scale independence'.

There is also some relation to the fractal concepts of Mandelbrot[87] which are based in notions of scale invariance. The 'partial' images generated at different levels by this model bear comparison with random fractal functions, and sections — consisting of the same number of pixels — from different levels of the tree would appear similarly 'rough' given similar β parameters. A texture field as generated by the model might be characterised by its β parameters by analogy with the fractal-based texture description scheme of Peleg et al[112] mentioned in section 2.2.

The realistic appearance of many of Mandelbrot's synthetic 'natural' scenes provides further evidence that image descriptions related to activity over a range of scales are potentially very valuable (see also [103],[113]).

2.4.3 Correlation Properties

The general correlation function for signals $f_{\cdot,\cdot,\cdot,\cdot}$ and $g_{\cdot,\cdot,\cdot,\cdot}$ is defined as

$$R_{fg}(l,i,j;k,p,q) = Ef_{l,i,j}g_{k,p,q} \quad . \tag{2.8}$$

Then from (2.4) and (2.5), the following three correlation functions may be defined:

(i) $R_{ww}(\cdot,\cdot,\cdot;\cdot,\cdot,\cdot)$

$$R_{ww}(l,i,j;k,p,q) = \delta_{l,k}\delta_{i,p}\delta_{j,q} \quad . \tag{2.9}$$

(ii) $R_{sw}(\cdot,\cdot,\cdot;\cdot,\cdot,\cdot)$

$$R_{sw}(l,i,j;r,m,n) = Es_{l,i,j}w_{r,m,n}$$

$$= \sum_{\substack{(k,p,q)\\ \in A_{l,i,j}}} \beta_k R_{ww}(k,p,q;r,m,n)$$

$$= \sum_{\substack{(k,p,q)\\ \in A_{l,i,j}}} \beta_k \delta_{k,r}\delta_{p,m}\delta_{q,n} \quad .$$

Thus

$$R_{sw}(l,i,j;r,m,n) = \begin{cases} \beta_r \ , & (r,m,n) \in A_{l,i,j} \ , \\ 0 \ , & (r,m,n) \notin A_{l,i,j} \ . \end{cases} \tag{2.10}$$

(iii) $R_{ss}(\cdot,\cdot,\cdot;\cdot,\cdot,\cdot)$

$$R_{ss}(l,i,j;r,m,n) = Es_{l,i,j}s_{r,m,n}$$

$$= \sum_{\substack{(k,p,q)\\ \in A_{l,i,j}}} \sum_{\substack{(d,a,b)\\ \in A_{r,m,n}}} \beta_k \beta_d E w_{k,p,q} w_{d,a,b}$$

$$= \sum_{\substack{(k,p,q)\\ \in A_{l,i,j}}} \sum_{\substack{(d,a,b)\\ \in A_{r,m,n}}} \beta_k \beta_d \delta_{k,d} \delta_{p,a} \delta_{q,b}$$

or

$$R_{ss}(l,i,j;r,m,n) = \sum_{\substack{(k,p,q)\in\\ (A_{l,i,j}\cap A_{r,m,n})}} \beta_k^2 \quad . \tag{2.11}$$

Note that the expected energy $Es^2_{l,i,j}$ increases monotonically with the level index l :

$$Es^2_{l,i,j} = R_{ss}(l,i,j;l,i,j)$$

$$= \beta_0^2 + \cdots + \beta_l^2 \quad . \tag{2.12}$$

It is possible following (2.11) to define a cumulative correlation function

$$B(l,i,j;k,p,q) = \sum_{\substack{(k,a,b)\\ \in D_{l,i,j}}} R_{ss}(k,a,b;k,p,q) \quad ,$$

$$k \geq l \quad . \tag{2.13}$$

This function represents the sum of the correlations between a node $s_{k,p,q}$ and all of the descendants of the node (l,i,j) which are on level k.

Now if $(k,p,q) \notin D_{l,i,j}$ then

$$R_{ss}(k,a,b;k,p,q) = R_{ss}(l,i,j;k,p,q) \quad ,$$

$$(k,a,b) \in D_{l,i,j} \quad , \tag{2.14}$$

and so

$$B(l,i,j;k,p,q) = 4^{k-l} R_{xx}(l,i,j;k,p,q) \quad . \tag{2.15}$$

Alternatively, if $(k,p,q) \in D_{l,i,j}$ then

$$B(l,i,j;k,p,q) = 4^{k-l} R_{xx}(l,i,j;l,i,j) + 4^{k-l-1}\beta_{l+1}^2 + \cdots + \beta_k^2 \tag{2.16}$$

or, by (2.12),

$$B(l,i,j;k,p,q) = 4^{k-l}(\beta_0^2 + \cdots + \beta_l^2) + 4^{k-l-1}\beta_{l+1}^2 + \cdots + \beta_k^2 \quad . \tag{2.17}$$

It is also possible to define a one-dimensional average correlation as a function of displacement parallel to one of the spatial coordinates on a level l:

$$T(l,p) = \frac{1}{N} \sum_{\substack{(l,m,n) \in D_{0,0,0}: \\ (l,m+p,n) \in D_{0,0,0}}} R_{xx}(l,m,n;l,m+p,n) \quad , \tag{2.18}$$

where N is the number of terms inside the summation.

Evaluation of (2.18) yields

$$T(l,p) = \sum_{k=0}^{l-1-n(p)} \beta_k^2 \left[\frac{1-2^{k-l}p}{1-2^{-l}p} \right] = \sum_{k=0}^{l-1-n(p)} b_k^l(p)\beta_k^2 \quad , \tag{2.19}$$

where $n(p)$ is defined by

$$2^{n(p)} \le p < 2^{n(p)+1} \tag{2.20}$$

and $n(0) = -1$.

The coefficient

$$b_k^l(p) = \frac{1 - 2^{k-l}p}{1 - 2^{-l}p} \tag{2.21}$$

of β_k^2 in (2.19) is monotonically nonincreasing with p.

By way of example, the average correlation between two nodes on level 8 of the quadtree at a separation of 30 nodes is given by

$$T(8,30) = \beta_0^2 + \frac{196}{226}\beta_1^2 + \frac{136}{226}\beta_2^2 + \frac{16}{226}\beta_3^2 \tag{2.22}$$

and that between nodes on level 5 at a separation of 7 is given by

$$T(5,7) = \beta_0^2 + \frac{18}{25}\beta_1^2 + \frac{4}{25}\beta_2^2 \ . \tag{2.23}$$

The function $T(l,p)$ may be seen from (2.19) to be a strictly nonincreasing function of p.

Taking the z transform of $T(l,p)$ in the variable p yields, assuming nonnegative definiteness of $T(l,p)$, a spectral density function (SDF) as

$$S(z) = \sum_{p=-(2^l-1)}^{2^l-1} \sum_{k=0}^{l-1-n(p)} \beta_k^2 \left[\frac{1 - 2^{k-l}p}{1 - 2^{-l}p} \right] z^{-p} \tag{2.24}$$

and frequency response on the unit circle in the z plane as

$$S(e^{j\omega}) = \beta_0^2 + \cdots + \beta_l^2 + 2\sum_{p=1}^{2^l-1}\sum_{k=0}^{l-1-n(p)}\beta_k^2\left[\frac{1-2^{k-l}p}{1-2^{-l}p}\right]\cos p\omega \quad . \qquad (2.25)$$

2.4.3.2 Long-Range Structure

Unlike most common image models such as the Markov random field (MRF) [153], the present model exhibits structure (i.e. correlation) over considerable distances in the image. This follows because any two image nodes with a given lowest common ancestor (LCA) have the same correlation as any other two nodes with the same LCA. One may therefore speak of image 'blocks' being correlated.

This is loosely analogous to the periodic multivariate random field model of Kashyap[72], in which blocks of an image are each described by a model and the relationship between these models is governed by a vector random field. In the present case each block has an 'internal' model — the martingale of (2.6) below some level l — and the relationship between blocks is governed by the same martingale above the level l .

2.4.3.3 Dyadic Shift Invariance

The correlation function of (2.11) is invariant to permutation of the spatial positions of the descendants of the nodes (l, i, j) or (r, m, n) within their respective descendant sets $D_{l,i,j}$ and $D_{r,m,n}$.

Such a permutation corresponds to a dyadic shift of blocks within the image, and hence the image correlation function is dyadic shift invariant[5].

2.4.4 Analogy with Fractal Surfaces

If a level l of the quadtree as generated by the model is visualised as a solid composed of rectangular blocks, each with height $s_{l,i,j}$ (which may be negative) and base area 4^{-l} then as l increases, the surface of this solid becomes increasingly subdivided and its area also increases.

As l becomes arbitrarily large, and assuming nonzero values of β_l, the surface area increases without bound while its projection on the horizontal plane (the total base area) remains fixed.

In this sense the surface is a random fractal[87] and must have a fractal dimension of between 2 and 3. In fact, since the process is gaussian and that density admits of arbitrarily large values (albeit with probability $P(\infty) \rightarrow 0$) the surface must, as $l \rightarrow \infty$, fill the entire three-dimensional volume above and below the base in the sense that any point in that volume is approached arbitrarily closely by the surface. Thus its fractal dimension must be 3.

Note that, like the area, the average energy of a node on the surface as given by (2.12) is unbounded as $l \rightarrow \infty$.

If, however, only a finite number of the β_l are nonzero then the surface has finite area and a fractal dimension of 2.

It seems likely that the intermediate case, where $\beta_l \rightarrow 0$ as $l \rightarrow \infty$, would give a fractal surface of dimension between 2 and 3, the exact value depending presumably on how fast β_l approaches zero.

2.4.5 Range of the Model

The image generated by the model possesses a block structure, with blocks of all scales superimposed. The parameters β_l determine the expected degree of activity at each scale. A few degenerate cases illustrate the range of the model.

- If $\beta_l = 0$ for $l > 0$, the model generates an image which consists of a uniform gray level.

- If $\beta_l = 0$ for $l < Y$, the model generates a gaussian white noise field of variance β_Y^2.

- If, for some k, $\beta_l = 0$ for $l \neq k$ and $\beta_k \neq 0$, the model generates a pattern of 'block' white noise at a scale which depends on k. The smaller the value of k, the larger is the block size of the generated pattern.

Thus the model represents the image as $(Y+1)$ superimposed 'block white noise' images, with scales ranging from the entire image to single pixels, with varying weights. The similar but separate treatment of different scales corresponds to the twin notions of scale invariance and scale independence alluded to in section 2.4.2.

There is a correspondence with 'progressive refinement' or 'progressive transmission' algorithms, as applied to image coding[73], in that as the quadtree is descended the image becomes increasingly complicated with smaller and smaller structure. The intermediate levels of the quadtree resemble the block-structured images used by Harmon[45] in his work on the recognition of faces.

2.5 Examples of Model-Generated Images

Figures P1 to P3 show examples of images synthesised by the model of (2.4) or (2.6). The upper photograph in each figure shows level Y of the quadtree and the lower shows levels 0 (at the bottom) to $Y-1$.

In figure P1, the model parameters β_l are equal, $\beta_l^2 = 1$. In figure P2, $\beta_l^2 = (3/2)^l$ and in figure P3, $\beta_l^2 = (2/3)^l$. The values of β_l^2 for each level l in each figure are given in table 2.1.

The same gaussian random number generator and seed were used for each of the three figures. The gray levels were scaled so that each image occupies the entire brightness range (black = 0, white = 255) of the framestore.

Figure β_l^2 =	P1 1	P2 $(3/2)^l$	P3 $(2/3)^l$
Level l		β_l^2	
1	1	1.500	0.6667
2	1	2.250	0.4444
3	1	3.375	0.2962
4	1	5.063	0.1975
5	1	7.594	0.1317
6	1	11.39	0.0878
7	1	17.09	0.0585
8	1	25.63	0.0390
9	1	38.44	0.0260

Table 2.1

Values of β_l^2 for Figures P1 — P3

2.6 Comments on the Model-Generated Examples

The block structure of the generated images is apparent from figures P1 to P3. The block edges coincide with boundaries in the quadtree. Since the same (pseudo-)random generator and seed (initial condition) were used for each of the three figures, the $w_{l,i,j}$ of (2.4) or (2.6) are identical in each figure and the differences between the figures are due solely to the values of β_l employed.

If the β parameters were spatially variant, $\beta = \beta_{l,i,j}$, one would expect the degree of structure at any given scale to vary across the image.

Obviously, many natural images do not possess such simple geometric structure and an inexpensive method for reducing the blocking effect will be introduced in Chapter 4.

However, natural images are often represented very poorly by such models as the Markov random field, which tend to generate textures with structure at a particular scale which depends on the model order.

In the examples of figures P1 to P3, it is apparent that structure exists at all scales and hence, for the larger scales, over sizeable distances in the image. In this sense the model captures an important property of natural images (see sections 1.5 and 2.4.2).

The 'vertical' causality of the model may be seen by following a feature from one of the upper quadtree levels (lower in the figures) through the larger and successively more complicated images to level Y, the image plane.

2.7 Digital Filtering Interpretation of the Model

Figure 2.3 depicts the recursive form of the model (2.6) as a time-varying recursive digital filter. The time direction corresponds to the level index l of the quadtree. Neglecting the time-varying multiplier, the filter has a pole at $z = 1$ and is thus marginally stable.

2.8 Correlation Transforms

Only the one-dimensional case will be considered for the sake of clarity.

The model in one dimension may be written in the spatially-invariant case as

$$s_{l+1,i'} = s_{l,i} + \beta_{l+1} w_{l+1,i'} \quad . \tag{2.26}$$

A correlation matrix may then be defined as

$$R_l(i,j) = E s_{l,i} s_{l,j} = R_{ss}(l,i;l,j) \quad . \tag{2.27}$$

Now since

$$R_{ss}(l+1,i';l+1,j') = R_{ss}(l,i;l,j) + \beta_{l+1}^2 \delta_{i',j'} \quad , \tag{2.28}$$

it follows that

$$R_{l+1} = C_1 \otimes R_l + \beta_{l+1}^2 I_{l+1} \tag{2.29}$$

where

$$C_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{2.30}$$

and I_l is the $(2^l \times 2^l)$ identity matrix and $\otimes$ denotes the Kronecker product[37].

2.8.1 Hadamard Transform

The Hadamard transform matrix[5] is given recursively by

$$H_{l+1} = H_1 \otimes H_l \quad , \tag{2.31}$$

where

$$H_1 = 2^{-1/2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad . \tag{2.32}$$

Theorem — the correlation matrix is diagonalised by the Hadamard transform.

Proof — it may be shown that if H_l diagonalises R_l , i.e.

$$H_l R_l H_l = \Delta_l \tag{2.33}$$

where Δ_l is diagonal, then H_{l+1} diagonalises R_{l+1} .

From (2.29) and (2.31),

$$H_{l+1}R_{l+1}H_{l+1} = (H_1 \otimes H_l)(C_1 \otimes R_l)(H_1 \otimes H_l) + \beta_{l+1}^2 I_{l+1} \quad . \tag{2.34}$$

By the *mixed product rule* [37] for Kronecker products, (2.34) may be written as

$$H_{l+1}R_{l+1}H_{l+1} = H_1 C_1 H_1 \otimes H_l R_l H_l + \beta_{l+1}^2 I_{l+1} \tag{2.35}$$

$$= 2\Gamma_{00} \otimes \Delta_l + \beta_{l+1}^2 I_{l+1} \tag{2.36}$$

$$= \Delta_{l+1} \quad , \tag{2.37}$$

where

$$\Gamma_{00} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad . \tag{2.38}$$

The initial condition is given by

$$H_1 R_1 H_1 = \Delta_1 = \begin{bmatrix} 2\beta_0^2 + \beta_1^2 & 0 \\ 0 & \beta_1^2 \end{bmatrix} \tag{2.39}$$

and the proof is complete.

2.8.2 Haar Transform

The Haar transform matrix[5] is given recursively by

$$T_{l+1} = \left| \begin{array}{c} C_0 \otimes T_l \\ \hline D_0 \otimes I_l \end{array} \right| \tag{2.40}$$

where

$$C_0 = 2^{-1/2} \begin{bmatrix} 1 & 1 \end{bmatrix} \tag{2.41}$$

and

$$D_0 = \begin{bmatrix} 1 & -1 \end{bmatrix} . \tag{2.42}$$

Theorem — the correlation matrix is diagonalised by the Haar transform.

Proof — it may be shown that if T_l diagonalises R_l , i.e.

$$T_l R_l T_l^T = \Delta_l \tag{2.43}$$

where Δ_l is diagonal, then T_{l+1} diagonalises R_{l+1} .

From (2.29) and (2.40), and noting that $2C_0^T C_0 = C_1$,

$$T_{l+1} R_{l+1} T_{l+1}^T = \left| \begin{array}{c} C_0 \otimes T_l \\ \hline D_0 \otimes I_l \end{array} \right| (2C_0^T C_0 \otimes R_l) \left(C_0^T \otimes T_l^T \mid D_0^T \otimes I_l \right) + \beta_{l+1}^2 T_{l+1} T_{l+1}^T \tag{2.44}$$

$$= \left[\begin{array}{c} 2C_0C_0^T C_0 \otimes T_l R_l \\ \hline 2D_0C_0^T C_0 \otimes R_l \end{array}\right] \left(C_0^T \otimes T_l^T \mid D_0^T \otimes I_l \right) + \beta_{l+1}^2 I_{l+1} \quad . \tag{2.45}$$

but $C_0C_0^T = (1)$ and $D_0C_0^T = C_0D_0^T = (0)$, so (2.45) may be written as

$$\bar{T}_{l+1}\bar{R}_{l+1}\bar{T}_{l+1}^T = \begin{bmatrix} 2T_lR_lT_l^T & 0_l \\ 0_l & 0_l \end{bmatrix} + \beta_{l+1}^2 I_{l+1} \tag{2.46}$$

$$= 2\Gamma_{00} \otimes \Delta_l + \beta_{l+1}^2 I_{l+1} \tag{2.47}$$

$$= \Delta_{l+1} \quad , \tag{2.48}$$

where again

$$\Gamma_{00} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad , \tag{2.49}$$

and where 0_l is a zero matrix of appropriate dimension.

The initial condition is given by

$$T_1R_1T_1 = \Delta_1 = \begin{bmatrix} 2\beta_0^2 + \beta_1^2 & 0 \\ 0 & \beta_1^2 \end{bmatrix} \tag{2.50}$$

and the proof is complete.

2.9 Implications of the Model for Estimation

As noted in Chapter 1, the estimator which is optimal under the criterion of minimum mean squared error (MMSE) reduces to a linear functional of the available data for processes which are normal (i.e. with jointly gaussian probability density functions). Since the model is gaussian, the MMSE estimator will be linear.

The form of the model of (2.4) suggests that a one-dimensional estimator which is causal in the vertical (l) index may be an appropriate scheme. The innovations representation (2.6) of the model further suggests a recursive form for the estimator by analogy with causal innovations representations of time series with increasing data support, which often yield recursive (Kalman) estimators. Such estimators are computationally simple but the price paid is the restriction to causality. In the present case, however, the image is noncausal and the (causal) estimator would operate in a direction 'normal' to the image plane. The use of a one-dimensional estimator would avoid the problem of two-dimensional spectral factorisation[27][28][29] which can yield NSHP factors of infinite spatial order.

The benefits of such a scheme would additionally include the ability to estimate information at different scales in a uniform, isomorphic way. Typical causal estimators have an order — equal to the order of the underlying signal model — which determines the scales at which they are effective. Structure at smaller scales is blurred (the estimator order being too large to follow fast fluctuations) and noise processes at larger scales pass unsmoothed (the estimator order being too small to affect slowly-varying signals). The causal direction is often all too obvious by virtue of blurring and streaking effects[109][127][157] produced at edges normal to it, as the estimator adjusts to the non-stationarity in the input at the edge.

2.10 Generalisations

2.10.1 Generalisation of the Model

The model of (2.6) expresses the image as a sample of a martingale process, causal in the level index l of the quadtree. An obvious generalisation of this model is the first-order Markov process[110], where the propagated term in (2.6) is multiplied by a spatially-variant coefficient $\alpha_{l+1,\,l',\,f'}$,

$$s_{l+1,\,l',\,f'} = \alpha_{l+1,\,l',\,f'} s_{l,\,i,\,j} + \beta_{l+1,\,l',\,f'} w_{l+1,\,l',\,f'} \quad ,$$

$$(l+1,\,l',\,f') \in D_{l,\,i,\,j} \quad ,$$

$$0 \le l \le Y-1 \quad . \tag{2.51}$$

This model may be constructed in such a way that at the image plane, the images generated by (2.51) and (2.6) are identical. This is accomplished by suitable choice of the β parameters. At higher levels in the quadtree, however, the 'partial' images are different.

A further generalisation may be introduced by the use of a more general l^{th}-order Markov model to represent the signal on level $l+1$,

$$s_{l+1,\,l',\,f'} = \sum_{\substack{(k,\,p,\,q) \\ \in A_{l,\,i,\,j}}} \alpha^{l+1,\,l',\,f'}_{k,\,p,\,q} s_{k,\,p,\,q} + \beta_{l+1,\,l',\,f'} w_{l+1,\,l',\,f'} \quad ,$$

$$(l+1,\,l',\,f') \in D_{l,\,i,\,j} \quad ,$$

$$0 \le l \le Y-1 \quad . \tag{2.52}$$

Again, this model may be rendered equivalent to the model of (2.4) or (2.6) in the sense

of the generated image by appropriate choice of the β parameters. This is apparent by consideration of the nonrecursive form (2.4) of the model. The 'partial' images on levels other than the image plane are different again from those produced by the original model or those produced by the model of (2.51).

2.10.2 Generalisation to N-Dimensional Signals

The model generalises to N dimensions if a '2^N-tree' structure is defined. A noncausal signal in this N-dimensional space may be obtained by the one-dimensional martingale model operating in a hyperspace of dimension N+1 and causal in the extra dimension. The N-dimensional signal space is then the projection of this (N+1)-dimensional structure on to the N-dimensional hyperplane, just as the 3-dimensional structure of the quad-tree projects on to the 2-dimensional image plane.

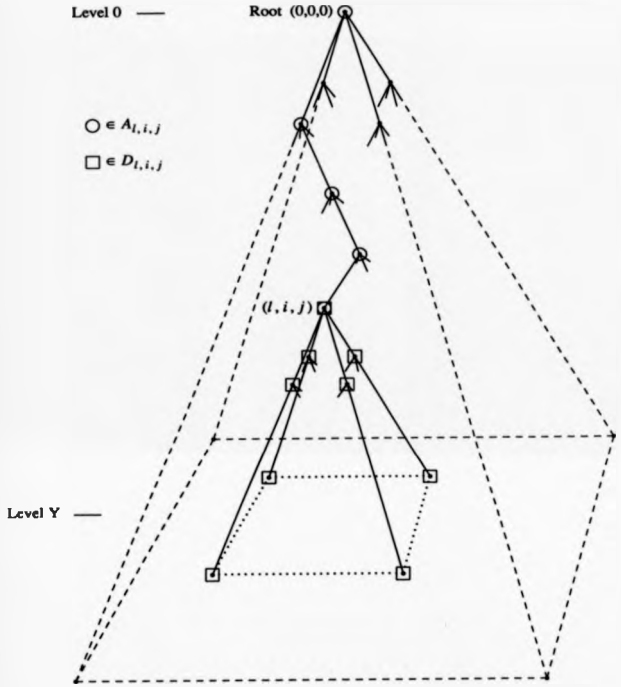


Figure 2.1 — Quadtree Structure

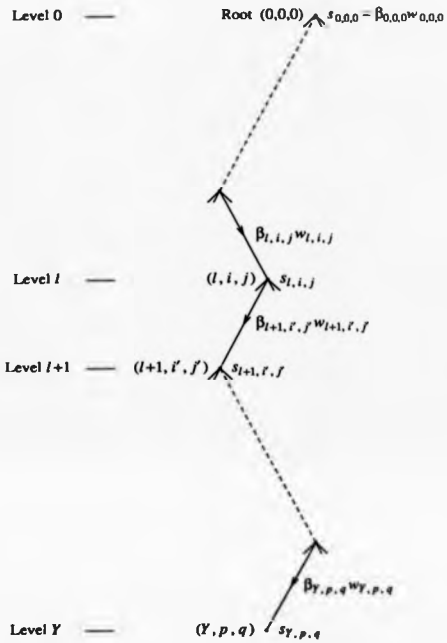


Figure 2.2 — Model Structure

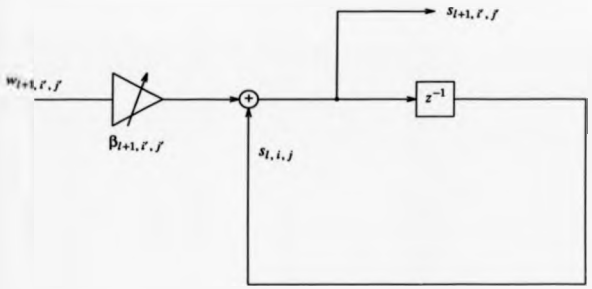


Figure 2.3 — Digital Filtering Interpretation of the Model

CHAPTER 3

THE ESTIMATOR

3.1 Problem Statement

Given noisy image data $u_{i,j}$, $0 \le i, j \le 2^Y - 1$, the general linear estimator (the optimal minimum mean squared error (MMSE) estimator was noted in Chapter 1 to be linear for normal processes) is given by

$$\hat{s}_{i,j} = \sum_{m,n} a_{m,n}^{i,j} u_{m,n} \quad . \tag{3.1}$$

In general, calculation of the estimate is computationally cumbersome, requiring in the worst case 4^Y multiply-accumulate operations per image pixel and both calculation of and allocation of storage for 4^Y coefficients $a_{m,n}^{i,j}$ per pixel. In any practical scheme, this burden would have to be substantially reduced, for example by assuming that correlations between pixels are significant only at short distances (the effect being that most of the $a_{m,n}^{i,j}$ are then zero) or that the statistics of the signal are stationary (implying that $a_{m,n}^{i,j} = a_{i-m,j-n}$). As noted in Chapters 1 and 2, however, each of these measures has adverse implications for the realism of the image model.

What is sought in this work (see section 2.9) is an optimal (i.e. MMSE) linear estimator with the benefits of the causal recursive structure of the model of Chapter 2 and, like the model, capable of scale-invariant and possibly spatially-variant (i.e. nonstationary) operation.

Thus it is desired to retain the advantages of MMSE estimation (tractability of the

optimal solution, linearity for normal processes, innovations representation leading to computationally simple causal recursive estimation) but in the context of a more realistic image model (noncausal, nonstationary, scale invariant) and at low computational cost (one-dimensional recursive implementation).

3.2 LMMSE Estimation and the Orthogonality Principle

If a (one-dimensional) signal $s(n)$ is estimated as $\hat{s}(n)$, two important definitions immediately result.

The estimation error $\varepsilon(n)$ is defined as

$$\varepsilon(n) = s(n) - \hat{s}(n) \tag{3.2}$$

and the mean squared error (MSE) as

$$P_n = E\varepsilon^2(n) = E[s(n) - \hat{s}(n)]^2 \quad . \tag{3.3}$$

It can be shown[7][99][110] that the MSE P_n is minimised when the estimate $\hat{s}(n)$ is the conditional expected value of the signal $s(n)$ given the available data $x(k)$,

$$\hat{s}(n) = E\{s(n) \mid x(k), k \in M_n\} \quad , \tag{3.4}$$

where M_n is the region of support of the available data for the estimate.

The discrete linear form of the estimator is given by

$$\hat{s}(n) = \sum_{k \in M_n} a_k^n x(k) \quad . \tag{3.5}$$

The solution for the optimal coefficients a_k^* is given by the *orthogonality principle* [110], which states that the estimation error $\varepsilon(n)$ should be orthogonal to all of the data $x(k)$:

$$E\,\varepsilon(n)x(k) = E[s(n) - \hat{s}(n)]\,x(k) = 0 \quad ,$$

$$k \in M_n \quad . \tag{3.6}$$

Application of the orthogonality relation for each datum generates the *normal* or *Yule-Walker* equations[86][110], solution of which yields the coefficients of the optimal estimator (see section 1.4).

3.3 Kalman Estimation and Innovations[68][86][110]

3.3.1 Causal Prediction on Increasing Data Support

Suppose that the task at hand is the linear prediction of the 'present' value $x(n)$ of a signal in terms of a linear combination of its causal past $x(k)$, $0 \le k \le n-1$. The predictor is denoted by $\hat{x}(n)$, and the mean squared prediction error is then given by

$$P_n = E[x(n) - \hat{x}(n)]^2 \quad . \tag{3.7}$$

This quantity is minimised if the predictor satisfies the orthogonality principle,

$$E[x(n) - \hat{x}(n)]\,x(m) = 0 \quad ,$$

$$0 \le m \le n-1 \quad . \tag{3.8}$$

The solution for the optimal predictor is simplified if $x(n)$ is expressed in terms of its *Kalman innovations* $i_x(n)$ [110]. This signal is orthonormal,

$$Ei_x(n)i_x(m) = \delta_{n,m} \quad . \tag{3.9}$$

and is derived from $x(n)$ by an invertible, causal linear transformation known as the Gram-Schmidt orthonormalisation[1][110]. This expresses $i_x(n)$ as a linear combination of $x(k)$ for $0 \leq k \leq n$,

$$i_x(n) = \sum_{k=0}^{n} \gamma_k^n x(k) \quad . \tag{3.10}$$

The inverse relation expresses $x(n)$ as a linear combination of $i_x(k)$ for $0 \leq k \leq n$,

$$x(n) = \sum_{k=0}^{n} l_k^n i_x(k) \quad . \tag{3.11}$$

The processes $x(n)$ and $i_x(n)$ are said to be *linearly equivalent* , since one may be derived from the other by the above causal linear transformations. The causal, time-varying linear filter with coefficients γ_k^n of (3.10) is known as the *Kalman whitening filter* [110] of $x(n)$, and that of (3.11) with coefficients l_k^n is known as the *Kalman innovations filter* [110] of $x(n)$.

From the relations (3.10) and (3.11) it is apparent that the estimate $\hat{x}(n)$ of $x(n)$ in terms of $x(k)$, $0 \leq k \leq n-1$ may be expressed as a linear combination of the innovations $i_x(k)$ for $0 \leq k \leq n-1$,

$$\hat{x}(n) = \sum_{k=0}^{n-1} \lambda_k^n i_x(k) \quad , \tag{3.12}$$

because these data are linearly equivalent to $x(k)$ for $0 \leq k \leq n-1$.

The orthogonality relation of (3.8) may then be written in terms of the innovations,

$$E[x(n) - \hat{x}(n)]\, i_x(m) = 0 \quad ,$$

$$0 \le m \le n-1 \quad . \tag{3.13}$$

It follows from (3.11) and (3.12) that (3.13) is satisfied if

$$\lambda_k^n = l_k^n \quad ,$$

$$0 \le k \le n-1 \quad , \tag{3.14}$$

and the optimal predictor is then given as

$$\hat{x}(n) = x(n) - l_n^n\, i_x(n) \quad . \tag{3.15}$$

Thus the signal $x(n)$ may be expressed as

$$x(n) = \hat{x}(n) + l_n^n\, i_x(n) \quad . \tag{3.16}$$

This description is known as a minimum variance representation (MVR) [86] of the signal $x(n)$ since it takes the form of a sum of the optimal predictor plus the minimum mean squared (i.e. minimum variance) error in the prediction.

The prediction error is

$$\varepsilon_n = x(n) - \hat{x}(n) = l_n^n\, i_x(n) \tag{3.17}$$

which is a nonstationary white noise, proportional to the Kalman innovations $i_x(n)$ with mean squared value

$$P_n = E\varepsilon_n^2 = (l_n^n)^2 \quad . \tag{3.18}$$

The innovations signal $i_x(n)$ may therefore be considered to represent that part of $x(n)$ which is 'unpredictable', and represents the 'new information' in $x(n)$ assuming knowledge of all previous $x(k)$, $0 \leq k < n$.

3.3.2 Kalman Estimation in White Noise

Section 3.3.1 describes the innovations approach to the solution of (3.5) by the orthogonality relation (3.6) for the case where $s(n) = x(n)$ and $M_n = \{k : 0 \leq k \leq n-1\}$. The estimator is the predictor of $x(n)$ in terms of its past.

Suppose now that it is desired to obtain an estimate $\hat{s}(n)$ of $s(n)$ in terms of $x(k)$ for $0 \leq k \leq n$ where the $x(k)$ are noisy observations of the signal $s(n)$, and the corrupting noise is white and orthogonal to the signal,

$$x(n) = s(n) + v(n) \quad , \tag{3.19}$$

$$E\,v(n)v(m) = Q_v(n)\delta_{n,m} \quad , \tag{3.20}$$

$$E\,s(n)v(m) = 0 \quad . \tag{3.21}$$

In the terms of (3.5), $M_n = \{k : 0 \leq k \leq n\}$. Since $s(n) \neq x(n)$, the problem is now one of filtering rather than of prediction. The data support M_n now includes the 'present' ($k = n$) as well as the 'past' ($0 \leq k \leq n-1$).

As in the case of the predictor of section 3.3.1, it will be shown that the derivation of the optimal (MMSE) estimator may be simplified if the estimator is expressed in terms of the Kalman innovations $i_x(n)$ of the data $x(n)$. It will further be shown that the estimator

may be coerced into a recursive form which simplifies its computation.

Lemma 1 [110]

The difference $x(n) - \bar{s}(n)$ between the present noisy observation $x(n)$ and the present MMSE estimate $\hat{s}(n)$ is proportional to the present value of the Kalman innovations $i_x(n)$:

$$x(n) - \bar{s}(n) = J_n i_x(n) \quad . \tag{3.22}$$

Proof

By the orthogonality relation of (3.6), for the estimate $\hat{s}(n)$ to be MMSE it follows that

$$E[s(n) - \hat{s}(n)]\, x(m) = 0 \quad ,$$

$$0 \le m \le n \quad . \tag{3.23}$$

Now from (3.19), (3.20) and (3.21), it follows that (3.23) may be rewritten in terms of $x(n) - \bar{s}(n)$ as

$$E[x(n) - \bar{s}(n)]\, x(m) = \underline{Q}_v(n)\delta_{n,m} \quad ,$$

$$0 \le m \le n \quad , \tag{3.24}$$

and so the difference $x(n) - \hat{s}(n)$ is orthogonal to $x(m)$ for $0 \le m \le n-1$,

$$E[x(n) - \hat{s}(n)]\, x(m) = 0 \quad ,$$

$$0 \le m \le n-1 \quad . \tag{3.25}$$

Now since the innovations signal $i_x(m)$ is a linear combination of the data $x(k)$ for $0 \le k \le m$, it follows that $x(n) - \bar{s}(n)$ must be orthogonal to the innovations $i_x(m)$ for $0 \le m \le n-1$,

$$E[x(n) - \bar{s}(n)]\, i_x(m) = 0 \quad ,$$

$$0 \le m \le n-1 \quad . \tag{3.26}$$

However, $x(n)$ and $\bar{s}(n)$ (and so also their difference) may be expressed as linear combinations of the innovations $i_x(m)$ for $0 \le m \le n$, and (3.22) follows from this fact and (3.26).

From (3.9) and (3.22), the coefficient J_n is given by

$$J_n^2 = E[x(n) - \hat{s}(n)]^2. \tag{3.27}$$

An important assumption is now made regarding the model for the signal $s(n)$. The signal will be assumed to be a martingale (c.f. the model of (2.6)) of the form

$$s(n+1) = s(n) + w(n+1) \quad , \tag{3.28}$$

$$E\, w(n) w(m) = Q_w(n) \delta_{n,m} \quad . \tag{3.29}$$

This model is a special case of a p^{th}-order Markov model (in this case $p = 1$).

Lemma 2

The difference $x(n+1) - \bar{s}(n)$ between the 'next' noisy observation $x(n+1)$ and the

present MMSE estimate $\hat{s}(n)$ is proportional to the 'next' value $i_x(n+1)$ of the Kalman innovations:

$$x(n+1) - \hat{s}(n) = H_{n+1} i_x(n+1) \quad . \tag{3.30}$$

Proof

The left hand side of (3.30) may be written as

$$x(n+1) - \hat{s}(n) = [s(n) - \hat{s}(n)] + w(n+1) + v(n+1) \quad . \tag{3.31}$$

The bracketed term on the right is orthogonal to $x(m)$ for $0 \leq m \leq n$ by (3.23), repeated here:

$$E[s(n) - \hat{s}(n)]\, x(m) = 0 \quad ,$$

$$0 \leq m \leq n \quad . \tag{3.23}$$

From (3.28) and (3.29) it follows that $w(n+1)$ is orthogonal to $x(m)$ over the same range of m,

$$E\, w(n+1)\, x(m) = 0 \quad ,$$

$$0 \leq m \leq n \quad , \tag{3.32}$$

and furthermore from (3.19) — (3.21), $v(n+1)$ is also orthogonal to $x(m)$ over this range,

$$E\, v(n+1)\, x(m) = 0 \quad ,$$

$$0 \leq m \leq n \quad . \tag{3.33}$$

Thus it can be seen that the left hand side of (3.30) or (3.31) is orthogonal to $x(m)$ for $0 \leq m \leq n$,

$$E[x(n+1) - \hat{s}(n)]\, x(m) = 0 \quad ,$$

$$0 \leq m \leq n \quad . \tag{3.34}$$

To prove Lemma 2, an argument similar to that of the proof of Lemma 1 above is invoked.

Since the Kalman innovations $i_x(m)$ are a linear combination of the data $x(k)$ for $0 \leq k \leq m$, it follows that the difference $x(n+1) - \hat{s}(n)$ is orthogonal to $i_x(m)$ for $0 \leq m \leq n$,

$$E[x(n+1) - \hat{s}(n)]\, i_x(m) = 0 \quad ,$$

$$0 \leq m \leq n \quad . \tag{3.35}$$

Now $x(n+1)$ may be expressed as a linear combination of the Kalman innovations $i_x(m)$ for $0 \leq m \leq n+1$. Also, $\hat{s}(n)$ may be expressed as a linear combination of the innovations $i_x(m)$ for $0 \leq m \leq n$. Thus the difference $x(n+1) - \hat{s}(n)$ may be written in terms of $i_x(m)$ for $0 \leq m \leq n+1$.

But since $x(n+1) - \hat{s}(n)$ is orthogonal to $i_x(m)$ for $0 \leq m \leq n$ by (3.35), then (3.30) follows and the proof is complete.

The value of H_{n+1} is given, from (3.9) and (3.30), as

$$H_{n+1}^2 = E[x(n+1) - \hat{s}(n)]^2 \quad . \tag{3.36}$$

It now remains to coerce the estimator into a recursive form.

Lemma 3

The 'next' estimate $\hat{s}(n+1)$ may be expressed in terms only of the 'present' estimate $\hat{s}(n)$ and the 'next' noisy observation $x(n+1)$,

$$\hat{s}(n+1) = a_{n+1}\,\hat{s}(n) + b_{n+1}\,x(n+1) \quad . \tag{3.37}$$

Proof

From the proof of Lemma 1 above,

$$\hat{s}(n+1) = x(n+1) - J_{n+1}\,i_x(n+1) \tag{3.38}$$

and, substituting (3.30) into (3.38),

$$\hat{s}(n+1) = x(n+1) - J_{n+1}\frac{[x(n+1) - \hat{s}(n)]}{H_{n+1}} \tag{3.39}$$

or

$$\hat{s}(n+1) = a_{n+1}\,\hat{s}(n) + b_{n+1}\,x(n+1) \quad , \tag{3.40}$$

where

$$a_{n+1} = \frac{J_{n+1}}{H_{n+1}} \tag{3.41}$$

and

$$b_{n+1} = 1 - a_{n+1} \quad . \tag{3.42}$$

which completes the proof.

The important consequence of this result is that the 'next' optimal (MMSE) estimate of the signal given all of the data may, for this class of data, be reduced to a combination of the 'present' estimate and the 'next' datum which involves only two multiplications per estimate rather than the linearly increasing operation count which would be required for the 'brute force' linear combination of all the available data.

3.4 Corruption by Additive White Gaussian Noise

In Chapter 2, the ('clean') image data is modelled as $s_{Y,i,j}$ as generated by the model of (2.6) where Y is the level index of the image plane in the quadtree.

In order to design the optimal estimator of the image given corrupted data, a model of the corrupting process is also required.

In this work, this process is modelled as additive white gaussian noise (AWGN) with zero mean and of known variance. This model was chosen because

(i) It frequently leads to mathematically highly tractable equations in the design of MMSE estimation schemes for smoothing, filtering and prediction of signals in noise;

(ii) When a number of independent sources of noise operate additively, their sum may be shown by the Central Limit Theorem[110] to possess asymptotically gaussian statistics, largely irrespective of the densities of the individual sources. The approximation is observed[110] to be rather good even for relatively few sources.

Real physical systems such as the electronic equipment used for capturing, processing, transmitting or displaying images tend to contain large numbers of active devices each of which contributes (generally non-white) noise. In addition, resistances (be they resistors by design, junction resistances in active devices, leakage resistances in reactive components or any other type) contribute white (Johnson) noise. To a first approximation, the combined effect of all these sources is their linear sum, thus admitting the Central Limit Theorem in its statistical description.

(iii) Additive white gaussian noise has been found in practice adequately to represent many of the forms of degradation observed in real physical systems which are (at least by design) linear. Nonlinear systems tend to give rise to various forms of multiplicative noise, the analysis of which is often difficult, but most systems of the type discussed in (ii) above are designed to be linear in their treatment of the substantive signal. This qualification is necessary because devices such as modulators and demodulators are nonlinear but their combination is usually intended to 'appear' linear or transparent.

3.4.1 The Model for the Noisy Image

Using the AWGN noise model as discussed above and the image model of Chapter 2, the combined model for the noisy observed image is given by

$$u_{i,j} = s_{Y,i,j} + v_{i,j} \tag{3.43}$$

where the probability density function (pdf) of v is gaussian and

$$Ev_{i,j}v_{k,l} = \sigma_v^2 \delta_{i,k}\delta_{j,l} \tag{3.44}$$

and

$$Es_{Y,i,j}v_{k,l} = 0 \tag{3.45}$$

It is assumed that the only available data is the noisy image $u_{i,j}$ and that the noise variance σ_v^2 is known or can be estimated.

3.5 Data Sets and Vertical Operations

The optimal (MMSE) estimator of (3.1) will be some type of generally nonstationary spatial linear operator. The estimate may be calculated at each image pixel by a multiply-accumulate operation at each point in the support (most generally the entire image) of that operator.

The computationally intensive nature of estimators which use some form of nonstationary spatial linear operator was mentioned in section 3.1.

In the present work, given the points above and the 'vertical' nature of the image model of (2.4) or (2.6), it is specifically desired that spatial operations be avoided. Accordingly, only vertical operations (i.e. those which are defined along the links in the quadtree structure) will be permitted. This prevents the estimator solution from emerging in the form of (3.1) although it may be expressed in that form as noted above. It will be shown that this

restriction does not adversely affect the optimality of the resulting estimator, because the constraint is matched to the structure of the model on which the estimator is based.

The implication is that all data flow in the estimation will be constrained to be along the links in the quadtree, and so at any node the 'data set' available for processing will be restricted to the ancestor and descendant sets for the given node.

The recursive form (2.6) of the model further suggests the development of a recursive form of the estimator based on the Kalman estimator which was developed in section 3.3.2. In this case, the data set might be further reduced (c.f. the remark at the end of section 3.3.2).

3.6 The Average-Value Data Quadtree

Given the noisy image $u_{i,j}$ as defined in (3.43), (3.44) and (3.45), a quadtree may be formed wherein each node (l, i, j) contains the average value of its four children $(l+1, i', j')$, $(l+1, i', j') \in D_{l,i,j}$ and the nodes on the image plane, level Y, are set equal to the values of the noisy data. The nodes of this tree will be denoted by $x_{\ldots}$.

Then

$$x_{Y,i,j} = u_{i,j} = s_{Y,i,j} + v_{i,j} \ ,$$

$$0 \le i, j \le 2^Y - 1 \ , \tag{3.46}$$

and

$$x_{l,i,j} = \frac{1}{4} \sum_{p=0}^{1} \sum_{q=0}^{1} x_{l+1,\,2i+p,\,2j+q} \ .$$

$$0 \le l \le Y-1 \ ,$$

$$0 \le i, j \le 2^l - 1 \ . \qquad (3.47)$$

It follows from (3.46) and (3.47) that the root node (0,0,0) contains the average value (gray level) of the entire image, while its four children each contain the average value of one quadrant of the image and so on.

This average-value quadtree will be of utility in the derivation of the optimal estimator in a subsequent section.

3.6.1 Correlation Properties

From the equations (2.1) to (2.4) and (3.43) to (3.47), and proceeding as in section 2.4.3, a number of correlation functions relating the data $x_{\ldots,\ldots}$, the signal $s_{\ldots,\ldots}$ and the signal innovations $w_{\ldots,\ldots}$ may be derived. The process is somewhat tedious, and so the derivations have been omitted; all the results follow, however, from the equations listed above.

(i) $R_{xw}(\ ,\ ,\ ,\ ;\ ,\ ,\ )$

$$R_{xw}(l,i,j;k,p,q) = Ex_{l,i,j}w_{k,p,q}$$

or

$$R_{xw}(l,i,j;k,p,q) = \begin{cases} \beta_k \ , & (k,p,q) \in A_{l,i,j} \ , \\ \dfrac{\beta_k}{4^{k-l}} \ , & (k,p,q) \in D_{l,i,j} \ , \\ 0 \ , & (k,p,q) \notin (A_{l,i,j} \cup D_{l,i,j}) \ . \end{cases} \qquad (3.48)$$

(ii) $R_{xx}(\ldots;\ldots)$

$$R_{xx}(l,i,j;k,p,q) = Ex_{l,i,j}x_{k,p,q}$$

$$= \begin{cases} R_{xx}(l,i,j;k,p,q)\ , & (k,p,q)\notin D_{l,i,j}\ , \\ R_{xx}(l,i,j;k,p,q)+\dfrac{\beta_{l+1}^2}{4}+\cdots+\dfrac{\beta_k^2}{4^{k-l}}\ , & (k,p,q)\in D_{l,i,j}\ . \end{cases}$$

Thus

$$R_{xx}(l,i,j;k,p,q) = \begin{cases} \displaystyle\sum_{\substack{(r,m,n)\in \\ (A_{l,i,j}\cap A_{k,p,q})}} \beta_r^2\ , & (k,p,q)\notin D_{l,i,j}\ , \\ \beta_0^2+\cdots+\beta_l^2+\dfrac{\beta_{l+1}^2}{4}+\cdots+\dfrac{\beta_k^2}{4^{k-l}}\ , & (k,p,q)\in D_{l,i,j}\ . \end{cases} \tag{3.49}$$

(iii) $R_{xx}(\ldots;\ldots)$

$$R_{xx}(l,i,j;k,p,q) = Ex_{l,i,j}x_{k,p,q}$$

$$= \begin{cases} R_{xx}(l,i,j;k,p,q)\ , & (k,p,q)\notin (A_{l,i,j}\cup D_{l,i,j})\ , \\ R_{xx}(l,i,j;k,p,q)+\dfrac{\beta_{l+1}^2}{4}+\cdots+\dfrac{\beta_\gamma^2}{4^{\gamma-l}}+\dfrac{\sigma_v^2}{4^{\gamma-l}}\ , & (k,p,q)\in D_{l,i,j}\ , \\ R_{xx}(l,i,j;k,p,q)+\dfrac{\beta_{k+1}^2}{4}+\cdots+\dfrac{\beta_\gamma^2}{4^{\gamma-k}}+\dfrac{\sigma_v^2}{4^{\gamma-k}}\ , & (k,p,q)\in A_{l,i,j}\ , \end{cases}$$

and so

$$R_{zz}(l,i,j;k,p,q) = \begin{cases} \sum_{\substack{(r,m,n)\in \\ (A_{l,i,j}\cap A_{k,p,q})}} \beta_r^2 , & (k,p,q)\notin(A_{l,i,j}\cup D_{l,i,j}) \\ \beta_0^2+\cdots+\beta_l^2+\dfrac{\beta_{l+1}^2}{4}+\cdots+\dfrac{\beta_Y^2}{4^{Y-l}}+\dfrac{\sigma_v^2}{4^{Y-l}} , & (k,p,q)\in D_{l,i,j} \\ \beta_0^2+\cdots+\beta_k^2+\dfrac{\beta_{k+1}^2}{4}+\cdots+\dfrac{\beta_Y^2}{4^{Y-k}}+\dfrac{\sigma_v^2}{4^{Y-k}} , & (k,p,q)\in A_{l,i,j} \end{cases} \tag{3.50}$$

3.6.2 The Optimal Upward Estimator

As a preliminary to the development of the optimal MMSE estimator, consider an estimator $\hat{s}^u_{l,i,j}$ of $s_{l,i,j}$ in terms only of data $x_{k,p,q}$ in the descendant set $D_{l,i,j}$ of the node (l,i,j),

$$\hat{s}^u_{l,i,j} = \sum_{\substack{(k,p,q)\\ \in D_{l,i,j}}} a^l_{k,p,q} x_{k,p,q} \tag{3.51}$$

It is apparent from the definitions (2.4), (2.6) of the model and (3.46), (3.47) of the average-value quadtree that:

(i) Since those nodes $(k,p,q)\in D_{l,i,j}$ contain averages of the nodes $(Y,p,q)\in D_{l,i,j}$ in the image plane, this estimate may be expressed in terms only of the latter nodes,

$$\hat{s}^u_{l,i,j} = \sum_{\substack{(Y,p,q)\\ \in D_{l,i,j}}} a^l_{Y,p,q} x_{Y,p,q} \tag{3.52}$$

and that

(ii) The $a^l_{Y,p,q}$ of (3.52) must be equal,

$$a^l_{Y,p,q} = a^l_Y \quad , \tag{3.53}$$

since the correlations between $s_{l,i,j}$ and all of the $x_{Y,p,q}$, $(Y,p,q) \in D_{l,i,j}$ are equal by (3.49).

Thus the estimate may be expressed in terms only of the average value $x_{l,i,j}$,

$$\hat{s}^u_{l,i,j} = a_l \, x_{l,i,j} \quad , \tag{3.54}$$

since all the image-plane descendants (Y,p,q) of the node (l,i,j) are weighted equally in the average.

By the orthogonality principle,

$$E[s_{l,i,j} - \hat{s}^u_{l,i,j}] x_{l,i,j} = 0 \quad , \tag{3.55}$$

giving for a_l

$$a_l = \frac{R_{xx}(l,i,j;l,i,j)}{R_{xx}(l,i,j;l,i,j)} \tag{3.56}$$

or

$$a_l = \frac{\beta_0^2 + \cdots + \beta_l^2}{\beta_0^2 + \cdots + \beta_l^2 + \frac{\beta_{l+1}^2}{4} + \cdots + \frac{\beta_Y^2}{4^{Y-l}} + \frac{\sigma_v^2}{4^{Y-l}}} \quad . \tag{3.57}$$

Thus $0 \leq a_l \leq 1$, and the LMMSE estimate $\hat{s}^u_{l,i,j}$ of $s_{l,i,j}$ in terms of all its descendants $x_{k,p,q}$, $(k,p,q) \in D_{l,i,j}$ in the average-value quadtree is just a positive fraction of the average $x_{l,i,j}$ at that node.

For estimation from a noiseless image ($\sigma_v^2 = 0$), (3.57) shows that the estimate is a larger positive fraction of $x_{l,i,j}$, which is now the average value of the 'clean' image block $s_{Y,p,q}$, $(Y,p,q) \in D_{l,i,j}$. Note that even in the noiseless case, $x_{l,i,j} \neq s_{l,i,j}$ in general for $0 \leq l \leq Y-1$.

3.7 The General Optimal Estimator

It is now desired to find the optimal MMSE estimator $\hat{s}_{l,i,j}$ in terms of all the data $x_{k,p,q}$ in the average-value quadtree. The restriction to vertical operations described in section 3.5 implies that only those values $x_{k,p,q}$, $(k,p,q) \in (A_{l,i,j} \cup D_{l,i,j})$ may be used in the estimate.

However, by an argument similar to that of section 3.6.2, this will still constitute the optimal estimator of $s_{l,i,j}$ in terms of *all* the $x_{Y,p,q}$ (i.e. all of the noisy data) since the correlations between the $s_{k,i,j}$ and the $x_{Y,p,q}$ are invariant within image blocks by (3.49) and so the contribution of a given block to the estimate may be expressed in terms only of the average value of the block.

Given that the optimal upward estimate $\hat{s}^u_{l,i,j}$ is simply a fraction of the average value $x_{l,i,j}$ (see section 3.6.2), it seems reasonable to enquire as to whether the optimal estimate in terms of $x_{k,p,q}$, $(k,p,q) \in (A_{l,i,j} \cup D_{l,i,j})$, may be expressed as a combination of $x_{k,p,q}$, $(k,p,q) \in A_{l,i,j}$ only. (Recall that $(l,i,j) \in A_{l,i,j}$ and so the average

value $x_{l,k_l,i}$ is included in the proposed set of available data for the estimate.)

As noted in sections 2.9 and 3.5, the innovations representation (2.6) of the model suggests that a recursive form of the estimator might be derived, which would achieve a further reduction in the size of the data set for the estimate.

The next few sections and appendices 1 and 2 address the derivation of the optimal MMSE estimator.

The recursive form of the estimator will be defined directly, in section 3.7.1, and shown to be optimal. Section 3.7.2 addresses the initial condition. Section 3.7.3 considers the estimation of the signal innovations process. A brief summary is given in section 3.7.4.

Appendix 1 derives the inverse of the data correlation matrix by recursion on the quadtree level index l and the nonrecursive form of the optimal estimator as the solution to the Yule-Walker equations. The recursive form of the estimator is then derived in a manner analogous to the algorithm of Levinson[84][110] which expresses the coefficients of the optimal predictor of order $N+1$ of a given process in terms of those of its optimal predictor of order N.

Appendix 2 presents a derivation of the estimator which is analogous to that given in section 3.3.2. In that case the noise was white and uncorrelated with the signal. In the present scheme such is not the case, but a similar recursive implementation of the estimator may nevertheless be obtained by using the innovations derived from the Kalman whitening filter of the data. Appendix 2 presents the whitening filter and the associated derivation of the recursive form of the estimator. The inverse of the data correlation matrix of Appendix 1 is shown to be given by the product of the Kalman whitening filter

and its matrix transpose. A similar decomposition of the correlation matrix as a product of the Kalman innovations filter and its matrix transpose is given. Both analyses constitute Cholesky factorisations[110] of their respective matrices into causal and anticausal factors.

3.7.1 Definition and Proof of Optimality

Theorem 1

The recursive estimator

$$\hat{s}_{l+1,\ell',f'} = \hat{s}_{l,i,j} + \beta_{l+1}\hat{w}_{l+1,\ell',f'} \quad ,$$

$$(l+1,\ell',f') \in D_{l,i,j} \quad , \tag{3.58}$$

of $s_{l+1,\ell',f'}$ is optimal in the MMSE sense with respect to all of the data $x_{\gamma,p,q}$, $(\gamma,p,q) \in D_{0,0,0}$, if $\hat{s}_{l,i,j}$ and $\hat{w}_{l+1,\ell',f'}$ are similarly optimal estimates of $s_{l,i,j}$ and $w_{l+1,\ell',f'}$ respectively.

Proof

The orthogonality relation for the MMSE optimality of the estimate $\hat{s}_{l+1,\ell',f'}$ may be written as

$$E\hat{s}_{l+1,\ell',f'}x_{\gamma,p,q} = Es_{l+1,\ell',f'}x_{\gamma,p,q} \quad ,$$

$$(\gamma,p,q) \in D_{0,0,0} \quad , \tag{3.59}$$

or

$$E\hat{s}_{l,i,j}x_{\gamma,p,q} + \beta_{l+1}E\hat{w}_{l+1,r,f}x_{\gamma,p,q} = Es_{l+1,r,f}x_{\gamma,p,q} ,$$

$$(\gamma,p,q)\in D_{0,0,0} . \tag{3.60}$$

As a consequence of the assumed optimality of $\hat{s}_{l,i,j}$ and $\hat{w}_{l+1,r,f}$, these estimators satisfy their respective orthogonality relations

$$E\tilde{s}_{l,i,j}x_{\gamma,p,q} = Es_{l,i,j}x_{\gamma,p,q} ,$$

$$(\gamma,p,q)\in D_{0,0,0} , \tag{3.61}$$

for $\tilde{s}_{l+1,i,j}$, and

$$E\tilde{w}_{l+1,r,f}x_{\gamma,p,q} = Ew_{l+1,r,f}x_{\gamma,p,q} ,$$

$$(\gamma,p,q)\in D_{0,0,0} , \tag{3.62}$$

for $\tilde{w}_{l+1,r,f}$.

Substituting these relations into the left side of (3.60), and then the model of (2.6) into both sides, gives (3.59) and the proof is complete.

In order to utilise the result (3.59), two further estimators are required.

The first is the optimal estimator $\hat{s}_{0,0,0}$ of the root node $s_{0,0,0}$ in terms of all of the data $x_{\gamma,p,q}$.

The second is the optimal estimator $\hat{w}_{l+1,r,f}$ of the signal innovations process $w_{l+1,r,f}$

in terms of all of the data $x_{Y,p,q}$.

3.7.2 Estimation of the Root Node

From section 3.6.2 it follows that the optimal MMSE estimator $\hat{s}_{0,0,0}$ of $s_{0,0,0}$ in terms of all of the data $x_{Y,p,q}$, $(Y,p,q) \in D_{0\,0\,0}$ is the upward estimate $\hat{s}^{\uparrow}_{0,0,0}$ of (3.54),

$$\hat{s}_{0,0,0} = \hat{s}^{\uparrow}_{0,0,0} = a_0 x_{0,0,0} \tag{3.63}$$

where, from (3.57), a_0 is given by

$$a_0 = \frac{\beta_0^2}{\beta_0^2 + \frac{\beta_1^2}{4} + \cdots + \frac{\beta_Y^2}{4^Y} + \frac{\sigma_v^2}{4^Y}} \; . \tag{3.64}$$

3.7.3 Estimation of the Innovations Process

The optimal MMSE estimator $\hat{w}_{l+1,f',f}$ of the signal innovations process $w_{l+1,f',f} = s_{l+1,f',f} - \hat{s}_{l,i,j}$ is required.

Theorem 2

The required MMSE innovations estimator is given by

$$\hat{w}_{l+1,f',f} = c_{l+1}(x_{l+1,f',f} - \hat{s}_{l,i,j}) \; ,$$

$$(l+1, f', f) \in D_{l,i,j} \; . \tag{3.65}$$

Proof

The orthogonality condition for the optimality of the estimate may be written as

$$E\hat{w}_{l+1,i',j'}x_{Y,p,q} = Ew_{l+1,i',j'}x_{Y,p,q} \quad ,$$

$$(Y,p,q)\in D_{0,0,0} \quad , \tag{3.66}$$

or

$$c_{l+1}\left(Ex_{l+1,i',j'}x_{Y,p,q} - E\bar{s}_{l,i,j}x_{Y,p,q}\right) = Ew_{l+1,i',j'}x_{Y,p,q} \quad ,$$

$$(Y,p,q)\in D_{0,0,0} \quad . \tag{3.67}$$

By the optimality of $\bar{s}_{l,i,j}$, it follows that the term $E\bar{s}_{l,i,j}x_{Y,p,q}$ may be replaced by $Es_{l,i,j}x_{Y,p,q}$.

Equation (3.67) may then be written in terms of the correlation functions introduced in section (3.6.1) as

$$c_{l+1}\left[R_{xx}(Y,p,q;l+1,i',j') - R_{xx}(Y,p,q;l,i,j)\right] = R_{xw}(Y,p,q;l+1,i',j') \quad ,$$

$$(Y,p,q)\in D_{0,0,0} \quad . \tag{3.68}$$

Considering first the case where $(Y,p,q)\in D_{l+1,i',j'}$, (3.68) yields

$$c_{l+1}\left[\beta_{l+1}^2 + \frac{\beta_{l+2}^2}{4} + \cdots + \frac{\beta_Y^2}{4^{Y-l-1}} + \frac{\sigma_v^2}{4^{Y-l-1}}\right] = \beta_{l+1} \quad , \tag{3.69}$$

and so (3.67) is satisfied for $(Y,p,q)\in D_{l+1,i',j'}$ if

$$c_{l+1} = \frac{\beta_{l+1}}{\beta_{l+1}^2 + \frac{\beta_{l+2}^2}{4} + \cdots + \frac{\beta_{\gamma}^2}{4^{\gamma-l-1}} + \frac{\sigma_v^2}{4^{\gamma-l-1}}} \quad . \tag{3.70}$$

Treating now the case where $(\gamma, p, q) \notin D_{l+1,i',f'}$, (3.68) yields

$$c_{l+1}\left[\sum_{\substack{(r,m,n)\in \\ (A_{l+1,i',f'} \cap A_{\gamma,p,q})}} \beta_r^2 - \sum_{\substack{(r,m,n)\in \\ (A_{l,i,j} \cap A_{\gamma,p,q})}} \beta_r^2\right] = 0 \quad . \tag{3.71}$$

But from the definition of the ancestor set,

$$A_{l+1,i',f'} = A_{l,i,j} \cup (l+1, i', f') \tag{3.72}$$

and

$$(l+1, i', f') \cap A_{\gamma,p,q} = \varnothing \tag{3.73}$$

where $\varnothing$ is the empty set.

Hence the parenthesised term in (3.71) is zero, and

$$(0)\, c_{l+1} = 0 \quad ,$$

$$(\gamma, p, q) \notin D_{l+1,i',f'} \quad , \tag{3.74}$$

to which the solution is arbitrary.

Hence the optimal MMSE innovations estimator is given by

$$\hat{w}_{l+1,i',f'} = c_{l+1}\,(x_{l+1,i',f'} - \hat{s}_{l,i,j}) \tag{3.75}$$

with c_{l+1} as given by (3.70).

Combining (3.70) and (3.65) gives

$$\hat{s}_{l+1,r,f} = \hat{s}_{l,i,j} + \left[\frac{\beta_{l+1}^2}{\beta_{l+1}^2 + \frac{\beta_{l+2}^2}{4} + \cdots + \frac{\beta_Y^2}{4^{Y-l-1}} + \frac{\sigma_v^2}{4^{Y-l-1}}} \right] (x_{l+1,r,f} - \hat{s}_{l,i,j}) \tag{3.76}$$

or

$$\hat{s}_{l+1,r,f} = e_{l+1}\hat{s}_{l,i,j} + (1 - e_{l+1})\, x_{l+1,r,f} \quad , \tag{3.77}$$

where

$$e_{l+1} = \frac{\frac{\beta_{l+2}^2}{4} + \cdots + \frac{\beta_Y^2}{4^{Y-l-1}} + \frac{\sigma_v^2}{4^{Y-l-1}}}{\beta_{l+1}^2 + \frac{\beta_{l+2}^2}{4} + \cdots + \frac{\beta_Y^2}{4^{Y-l-1}} + \frac{\sigma_v^2}{4^{Y-l-1}}} \tag{3.78}$$

or

$$e_l = \frac{\frac{\beta_{l+1}^2}{4} + \cdots + \frac{\beta_Y^2}{4^{Y-l}} + \frac{\sigma_v^2}{4^{Y-l}}}{\beta_l^2 + \frac{\beta_{l+1}^2}{4} + \cdots + \frac{\beta_Y^2}{4^{Y-l}} + \frac{\sigma_v^2}{4^{Y-l}}} \quad . \tag{3.79}$$

It is apparent from (3.77) and (3.78) that the optimal MMSE estimate $\hat{s}_{l+1,r,f}$ of a node $s_{l+1,r,f}$ is a weighted average of the optimal estimate $\hat{s}_{l,i,j}$ at its father and the (noisy) average value $x_{l+1,r,f}$ at the node itself. As the variance σ_v^2 of the corrupting noise increases, the estimator coefficient e_{l+1} increases toward unity. This implies that there

will be a greater contribution from the father node (i.e. more smoothing) and a lesser contribution from the noisy average at the current node as the noise variance increases.

The estimator coefficient at the bottom level (the image plane, level Y) of the quadtree is given by (3.79) as

$$e_Y = \frac{\sigma_v^2}{\beta_Y^2 + \sigma_v^2} \tag{3.80}$$

and so if the data is noiseless ($\sigma_v^2 = 0$) then the coefficient e_Y is zero, the estimate $\hat{s}_{Y,i,j}$ is equal to the data $x_{Y,i,j} = s_{Y,i,j}$ and the error in the estimation of the image plane is zero, as would be expected, irrespective of values of β_l at higher levels of the quadtree.

3.7.4 Summary

It has been shown that for the image model of Chapter 2 as given by (2.4) or (2.6), the recursive estimator of (3.77) and (3.78), defined on the quadtree structure and causal in the level index, is optimal in the minimum mean squared error (MMSE) sense with respect to all of the noisy image data.

The estimator equations are reproduced below; note that $(l+1, i', j') \in D_{l,i,j}$.

The estimator is given by the recursion

$$\hat{s}_{l+1,i',j'} = \hat{s}_{l,i,j} + \beta_{l+1}\hat{w}_{l+1,i',j'} \tag{3.58}$$

with initial condition (estimate at the root node)

$$\hat{s}_{0,0,0} = a_0 x_{0,0,0} \tag{3.63}$$

and innovations estimate

$$\hat{w}_{l+1,r,f} = c_{l+1}(x_{l+1,r,f} - \hat{s}_{l,i,j}) \quad . \tag{3.65}$$

Combining (3.58) and (3.65) gives a more directly useful form of the estimator,

$$\hat{s}_{l+1,r,f} = e_{l+1}\hat{s}_{l,i,j} + (1 - e_{l+1}) x_{l+1,r,f} \quad . \tag{3.77}$$

In the above equations, the various coefficients are given in terms of the parameters β_l of the image model as

$$a_0 = \frac{\beta_0^2}{\beta_0^2 + \frac{\beta_1^2}{4} + \cdots + \frac{\beta_Y^2}{4^Y} + \frac{\sigma_v^2}{4^Y}} \quad , \tag{3.64}$$

$$c_l = \frac{\beta_l}{\beta_l^2 + \frac{\beta_{l+1}^2}{4} + \cdots + \frac{\beta_Y^2}{4^{Y-l}} + \frac{\sigma_v^2}{4^{Y-l}}} \quad . \tag{3.70}$$

$$e_l = \frac{\frac{\beta_{l+1}^2}{4} + \cdots + \frac{\beta_Y^2}{4^{Y-l}} + \frac{\sigma_v^2}{4^{Y-l}}}{\beta_l^2 + \frac{\beta_{l+1}^2}{4} + \cdots + \frac{\beta_Y^2}{4^{Y-l}} + \frac{\sigma_v^2}{4^{Y-l}}} \quad . \tag{3.79}$$

3.7.5 Digital Filtering Interpretation of the Estimator

The recursive estimator of (3.77) may be represented as a time-varying digital filter as shown in figure 3.1. The time variable for the filter corresponds to the level index l for

the quadtree. The filter is clearly causal and recursive.

Neglecting the time variation of the multipliers, the filter effectively has a single pole on the real axis in the z-plane at $z = e_l$. Since $0 \le e_l \le 1$, the filter is stable in the bounded-input, bounded-output (BIBO) sense.

Comparison with figure 2.3 shows that the estimator and the underlying image model have similar structures as might be expected.

3.8 Computational Burden

The recursive form (3.77) of the estimator requires on average 5/4 = 1.25 multiplications per node to calculate $\bar{s}_{l+1,i',j'}$, since the first product $e_{l+1}\,\bar{s}_{l,i,j}$ needs to be computed only once for all four of the children $(l+1, i', j')$ and the second term $(1 - e_{l+1})\,x_{l+1,i',j'}$ requires one multiplication per child.

Since there are approximately $4N/3$ nodes in the quadtree for an image of N pixels, the average multiplication count per image pixel is

$$\frac{4}{3} \times \frac{5}{4} = \frac{5}{3} \quad . \tag{3.81}$$

The formation of the average-value quadtree requires only additions and trivial multiplications by 1/4, which may be implemented by a double right-shift in a fixed-point binary format or by a double decrement of the exponent in a floating-point format.

The combined operation count of 5/3 multiplications per pixel compares very favourably indeed with that of other image estimation schemes and represents a reduction over a

(15×15) element spatial operator, as used by the authors of [75], of a factor of 135, or more than two orders of magnitude. Over even a small (5×5) spatial operator, the acceleration is by a factor of 15.

These figures do not include the overheads involved in other tasks such as the computation of the estimator coefficients (one per level of the quadtree in the spatially invariant case) or addressing.

The vertical structure of the algorithm renders it highly appropriate for parallel computation by an array of processing elements, which is not true of common causal image estimation schemes where the causality is in, rather than 'normal' to, the image plane. On the other hand, conventional noncausal estimators, whilst amenable to parallel computation, tend to be computationally expensive as the examples above demonstrate.

Thus the present scheme both eliminates the sequential processing demanded by traditional causal methods, and dramatically reduces the computational load associated with noncausal techniques. The computational burden in terms of operation count per pixel is significantly lower than that of even a fairly low-order causal scheme.

3.9 Estimation of the Estimator Coefficient

The optimal estimator coefficient e_l for level l of the quadtree is given by (3.79) in terms of the model parameters β_l^2.

Since no prior knowledge of these values is assumed, the estimator coefficient must itself be estimated from the noisy image data $x_{k,p,q}$ in the average-value quadtree.

Now (3.79) may be rewritten as

$$e_l = \frac{N_{l+1}}{4N_l} \ ,$$

$$1 \le l \le Y-1 \ , \tag{3.82}$$

where

$$N_{k+1} = 3\left[\frac{\beta_{k+1}^2}{4} + \cdots + \frac{\beta_Y^2}{4^{Y-k}} + \frac{\sigma_v^2}{4^{Y-k}}\right] \ . \tag{3.83}$$

From (3.50) it follows that the N_k may be written in terms of

$$F_k = E x_{k,p,q}^2 \tag{3.84}$$

as

$$N_{k+1} = R_{xx}(k+1, m, n; k+1, m, n) - R_{xx}(k, a, b; k, a, b) \ ,$$

$$= F_{k+1} - F_k \ ,$$

$$0 \le k \le Y-1 \ , \tag{3.85}$$

which is just the difference between the expected energies of nodes on levels $k+1$ and k of the average-value quadtree.

If N_{Y+1} is defined as

$$N_{Y+1} = 3\,\sigma_v^2 \tag{3.86}$$

then (3.82) holds additionally for $l = Y$.

The statistics N_k are now required in order to calculate the optimal estimator coefficients. These expected values are not available and must be estimated from the noisy data.

The obvious and consistent estimator of the expected energy $F_l = R_{xx}(l, i, j; l, i, j)$ is simply the sample average

$$\hat{F}_l = 4^{-l} \sum_{\substack{(l,i,j) \\ \in D_{0,0,0}}} x^2_{l,i,j} \tag{3.87}$$

of the energies of all the nodes on level l of the quadtree.

Then

$$\hat{N}_{l+1} = \hat{F}_{l+1} - \hat{F}_l \quad ,$$

$$0 \le l \le Y-1 \quad , \tag{3.88}$$

is the estimator of N_{l+1}, and

$$\hat{e}_l = \frac{\hat{N}_{l+1}}{4\hat{N}_l} \quad ,$$

$$1 \le l \le Y-1 \quad , \tag{3.89}$$

is the estimator of e_l for $1 \le l \le Y-1$.

The estimator of e_Y is given by

$$\hat{e}_Y = \frac{N_{Y+1}}{4\hat{N}_Y} \tag{3.90}$$

with N_{l+1} as defined by (3.86).

3.9.1 The Probability Density Function for the Estimator Coefficient

The probability density function (pdf) of the (estimated) optimal estimator coefficient $\hat{e}_l$ is a function of those of the $\hat{N}_m$ of (3.89).

These quantities may be expressed in terms of the differences between father and child nodes since

$$\left[\frac{1}{4}\sum_{\substack{(l+1,\ell',f')\\ \in D_{l,i,j}}} x^2_{l+1,\ell',f'}\right] - x^2_{l,i,j} = \frac{1}{4}\sum_{\substack{(l+1,\ell',f')\\ \in D_{l,i,j}}} m^2_{l+1,\ell',f'} \quad . \tag{3.91}$$

where

$$m_{l+1,\ell',f'} = x_{l+1,\ell',f'} - x_{l,i,j} \tag{3.92}$$

is the difference between the values of a child and its father, and the $\hat{N}_k$ are spatial averages of the squares of these differences by (3.87) and (3.88).

It follows from (3.46), (3.47) and (3.92) that the difference $m_{l+1,\ell',f'}$ is a gaussian variable since it is a linear combination of other gaussian quantities. Its mean is zero and its variance will be denoted $\sigma^2_{m_{l+1}}$.

Then assuming independence of these terms, the pdf of $\hat{N}_{l+1}$ is a chi-square (χ^2) density[8] with 4^{l+1} degrees of freedom and variance parameter $\sigma^2_{m_{l+1}}$. This follows from (3.87), (3.88), (3.91) and (3.92) which together allow $\hat{N}_{l+1}$ to be expressed as an average

of 4^{l+1} squares of gaussian variables $m_{l+1,i,j}$ each of which has variance $\sigma^2_{m_{l+1}}$.

Similarly, $\tilde{N}_l$ has a χ^2 density with 4^l degrees of freedom and variance parameter $\sigma^2_{m_l}$.

Since $\tilde{e}_l$ is the ratio (3.89) of these two χ^2 variables, its density is the so-called F density[8] which is characteristic of a ratio of χ^2 variables. Appendix 3 provides a full derivation.

The variance of $\tilde{e}_l$ increases with decreasing quadtree level l (i.e. as the pyramid is ascended) as the sample size (the number of nodes on the level l) is reduced.

Fortunately the problem is well conditioned in that the $\tilde{e}_l$ on the upper (small l) levels, while prone to larger errors in estimation, have a correspondingly smaller effect on the final image estimate, as is evident from the recursion of (3.77).

The estimate $\tilde{e}_l$ of e_l is consistent in that as l increases, $\tilde{e}_l \underset{l\to\infty}{\to} e_l$ as the variances of the $\tilde{N}_m$ approach zero.

3.10 Examples of Estimated Images

The estimator was applied to noisy versions of the 'girl' and 'blobs' images, the clean originals of which appear in figures P4 and P5. The images were processed at an input signal-to-noise ratio of 0dB. The noisy images are shown in quadtree form in figures P7 and P8, and the estimated images in figures P9 and P10, as indicated in table 3.1. All estimated images displayed in this work are scaled to the same variance as the original uncorrupted image.

Image	'Girl'	'Blobs'
Clean	P4	P5
Noisy (0dB)	P7	P8
Estimated	P9	P10
SNR of estimate	10.9dB	11.4dB

Table 3.1

Estimation Results

3.10.1 Comments on the Estimated Examples

It is clear from figures P9 and P10, and from the SNR figures given in table 3.1, that the estimator achieves some degree of noise reduction. However, the immediate subjective impression is that of the pronounced block structure of the estimated images. This is a consequence of:

(i) the quadtree structure and the lateral isolation between nodes which it provides;

(ii) the use of the simple averaging procedure of (3.47) which gives rise to alias distortion when the data on level $l+1$ is averaged and the spatial sampling rate reduced in the calculation of the data on level l;

(iii) the form (3.77) of the downward estimation process which, because of the definition of the model and estimator on the quadtree, maintains the isolation referred to in (i) above.

There is also some blurring of the image; this is due to the use of the spatially-invariant form of the estimator, which results in an inability adequately to handle nonstationarities (i.e. edges) as discussed in section 2.1.

These problems will be addressed in Chapter 4, which introduces both an inexpensive method of reducing the blocking effect and also the spatially-variant form of the estimator.

The SNR improvement (10.9dB and 11.4dB for the 'girl' and 'blobs' images respectively at input SNR of 0dB) is very reasonable, however, and represents performance numerically superior to that of many existing restoration schemes[61][156]. This suggests that even the spatially-invariant form of the quadtree image model is a fairly effective statistical description of the two images under consideration.

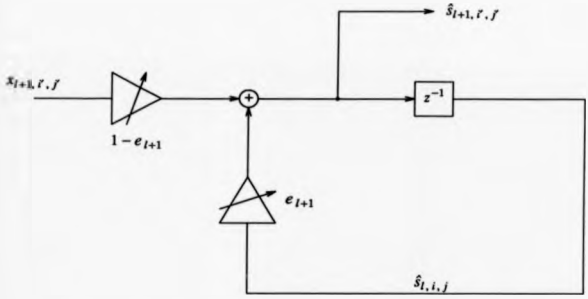


Figure 3.1 — Digital Filtering Interpretation of the Estimator

CHAPTER 4

ADAPTATION OF THE ESTIMATOR TO VISUAL CRITERIA

4.1 Motivation

As noted at the end of Chapter 3, the estimator as it stands suffers from two key defects.

The first is evident in the blocking effects in the estimated images, which are a consequence of the structure of the quadtree and the nature of the upward averaging and downward estimation processes (involved in calculation of the data and estimated quadtrees respectively). Other than the upward averaging process, there is no explicit lowpass ('anti-alias') filtering of the data prior to reduction of the sample density of the sort which is commonly designed into sampled-data systems because of this effect. Indeed, this kind of spatial operator has been deliberately avoided in order to reduce the computational complexity of the scheme. The averaging process constitutes a zero-order hold (ZOH) lowpass operator which possesses sidelobes of significant energy in the frequency domain and hence is a rather poor lowpass filter. The downward estimation, which is where the blocking effects become apparent, suffers again as a result of the isolation between nodes which do not share immediate ancestors.

It will be shown in this chapter that the problem may be largely solved, and the blocking effect dramatically reduced, by the inexpensive expedient of inserting relatively few additional nodes into the quadtree structure. Spatial filtering is not required, and the computational burden is only marginally increased.

Secondly, the use of the spatially-invariant form of the model and estimator leads to

blurring of salient image features such as lines and edges, which are profoundly space-variant features. The estimator performs the same operations at every node on any given level of the quadtree and so cannot adjust the degree of smoothing applied in different regions of the image.

This can be considered as a combination of defects in both the spatially-invariant signal model and the MMSE error criterion (see section 1.6). However, these effects are not easily isolated and the design of a globally optimal estimator for a signal model and error criterion which are both significantly more complicated is likely to prove difficult as noted in section 1.6. Instead, a more pragmatic approach will be followed; adaptations will be introduced which preserve the efficient tree-based structure of the estimator.

This chapter will introduce the spatially-variant or nonstationary form of the estimator. Contextual information concerning the local structure of the image will be extracted by a computationally efficient 'edge detector' (actually a more general scale-invariant *activity* detector) which operates on the quadtree. This information will be used to modify the degree of smoothing which the estimator effects in the locality and at the given scale.

By a process of backward 'extrapolation', the corresponding spatially-variant form of the image model may be derived. This is effectively a 'signal-equivalent' technique analogous to those of Abramatic and Silverman[3] and Knutsson, Wilson and Granlund[75].

The combined system will be shown to offer the benefits of scale-invariant, noncausal and nonstationary estimation but at very low computational cost.

4.2 Reduction of Blocking and Alias Distortion

The structure of the quadtree, and the definitions (3.47) and (3.77) of the upward averaging and downward estimation processes respectively, give rise to noticeable block edge artifacts in the estimated image.

Nothing can be done about this effect without sacrificing to some extent the strictly vertical nature of the algorithm.

However, in the present scheme it has been found that a dramatic reduction of the scale of this problem can be achieved with little such sacrifice by the inexpensive expedient of inserting some extra nodes into intermediate positions in the quadtree structure.

Figure 4.1 shows the locations of these extra nodes, which will be referred to as *interstitial* nodes. The existing nodes are known as *original* nodes.

On a level l, $1 \le l \le Y-1$ in the quadtree there are $(2^l)^2$ original nodes $s_{l,i,j}$, $0 \le i, j \le 2^l - 1$. Between these nodes are inserted $(2^l-1)^2$ interstitial nodes $p_{l,m,n}$, $0 \le m, n \le 2^l - 2$, such that each $p_{l,m,n}$ lies centrally between the four original nodes $s_{l,m+s,n+t}$, $s, t \in \{0, 1\}$.

The children of an interstitial node $p_{l,m,n}$ are then the four original nodes $s_{l+1,2m+1+s,2n+1+t}$, $s, t \in \{0, 1\}$, which lie closest to its downward projection on to level $l+1$. The interstitial node is referred to as the *interstitial father* of each of these four children; each child also has an original father, the four respective original fathers being the four original nodes on level l between which the interstitial node was inserted.

Figure 4.1 demonstrates that the interstitial node $p_{l,m,n}$ straddles a pair of block edges on level $l+1$.

Note that if the unmodified quadtree contained N nodes, then there are approximately $N/4$ new interstitial nodes, which together straddle all the block boundaries in the tree.

The computation is modified as follows:

(i) When forming the average-value quadtree, the value $y_{\ldots\ldots}$ of an interstitial node is defined simply as the average of its four (original) children, just as in the case of an original node:

$$y_{l,i,j} = \frac{1}{4}\sum_{m=1}^{2}\sum_{n=1}^{2} x_{l+1,2i+m,2j+n} \quad ,$$

$$1 \le l \le Y-1 \quad ,$$

$$0 \le i,j \le 2^l - 2 \quad . \tag{4.1}$$

(ii) The downward estimation as given by (3.77) is modified such that the information propagated downward is now the average of estimates at the original and interstitial fathers of the node being estimated,

$$\hat{s}_{l+1,r,f} = e_{l+1}\left[\frac{\hat{s}_{l,i,j} + \hat{p}_{l,s,t}}{2}\right] + (1-e_{l+1})\,x_{l+1,r,f} \tag{4.2}$$

where $p_{l,s,t}$ is the interstitial father of $s_{l+1,r,f}$.

(iii) The interstitial nodes are estimated as

$$\hat{p}_{l,i,j} = e_l \hat{g}_{l,i,j} + (1-e_l) y_{l,i,j} \tag{4.3}$$

where $\hat{g}_{l,i,j}$ is an average of the estimates of the four original nodes which surround the interstitial node $p_{l,i,j}$. The value $\hat{g}_{l,i,j}$ may be considered as an estimate of a hypothetical father of the interstitial node, which does not have a father in the sense that an original node does and so one must be created.

There are effectively two interdependent estimation problems, one concerning the original and the other the interstitial nodes. The approach adopted is suboptimal in that it addresses each independently, and in that no attempt is made to derive the globally optimal solution, which would lack the computationally efficient recursive form since the interstitial node structure is nonrecursive.

The insertion of the interstitial nodes reduces the blocking effect by partially destroying the complete isolation which exists between the descendants of neighbouring nodes on any given level of the unmodified quadtree.

In the context of the upward averaging process, the modification may be expressed in terms of the theory relating to the reduction in the two-dimensional spatial sampling density which occurs as the tree is ascended. The averaging method of forming the data quadtree, while computationally simple, does not involve any other lowpass filtering before the data is resampled at the next level at a lower rate. This resampling corresponds to a two-dimensional *decimation* [120][142] and, without prior lowpass filtering, induces frequency components in the downsampled representation which were not present at the higher sampling rate. This is the phenomenon of *aliasing* [120][142] familiar from discrete-time signal processing theory.

Watson[142] and Meer, Bauer and Rosenfeld[92] have considered the frequency-domain properties of pyramid-generating kernels. Both papers point out that the averaging method will lead to aliasing since the kernel has a $(sinx)/x$ form in the frequency domain. However, the present modification to the quadtree structure cheaply eliminates the filtering which would be required in both the upward averaging and downward estimation stages. A scheme such as the Laplacian pyramid of Burt and Adelson [11] which uses lowpass filtering prior to the decimation will not significantly exhibit the phenomenon of aliasing, but will be accordingly of increased computational complexity and will necessarily involve spatial operations. If the structure were to be used for estimation, spatial filtering (interpolation) would be required at the downward estimation stage as well.

The addition of the interstitial nodes in the present scheme and the geometry of the modified structure cause a cancellation (to a first-order approximation) of the alias components. This is so because the interstitial nodes on a given level form a lattice which is displaced from the lattice of original nodes by half of the sampling interval in each spatial coordinate. An alias is generated by the averaging of (4.1) in the same fashion as in the unmodified structure, but at the opposite phase for components of the signal above half the new sampling frequency. The use of the average of (4.2) in the downward estimation causes these components to cancel with those generated by the lattice of original nodes, and hence the aliasing is eliminated without recourse to conventional spatial filtering.

Note that this method does not achieve total cancellation because the only linear directions in which the sampling density is doubled by the addition of the interstitial nodes are the two diagonals. In the unmodified scheme, the number of nodes decreases by a factor of four from one level to the next, whereas in the modified scheme this factor is reduced

effectively to two. (The 'missing' factor of 2 lies in the fact that when calculating the next level, only data at the original nodes on the current level are used. There is still a factor of 4 fewer nodes at each successive higher level). A factor of unity would be required in theory for totally alias-free operation in the absence of an ideal lowpass filter.

The modified form of the estimator is derived heuristically as noted above and is not optimal in the MMSE sense for the given signal model. Nonetheless, it achieves a significant improvement of the estimate both subjectively and numerically, and may be regarded as a response to a defect in the original signal model.

4.3 Examples of Images Estimated with Interstitial Nodes

Examples of images estimated with the modification of section 4.2 are shown in figures P11 and P12. The noisy input images are the same as were used in the examples of section 3.10, the only difference being the inclusion of the interstitial nodes in the present cases.

Table 4.1 provides a key to the figures and, for comparison, to the figures presented in section 3.10. The present estimates are referred to in the table as 'modified' estimates.

Image	'Girl'	'Blobs'
Clean	P4	P5
Noisy (0dB)	P7	P8
Estimated (sec. 3.10)	P9	P10
Modified estimate	P11	P12
SNR (sec. 3.10)	10.9dB	11.4dB
SNR (modified)	12.6dB	13.0dB

Table 4.1

Estimation Results with Interstitial Nodes

It is clear from the figures that the insertion of the interstitial nodes and the modification of the estimation scheme as detailed in section 4.2 have completely removed (at least to below the threshold of visibility) the problem of blocking which was so evident in the unmodified estimates.

The further improvement of about 1.6dB in the SNR of the restorations by comparison with the unmodified scheme of section 3.10 demonstrates that the modification, while departing from the optimality of the method, does nevertheless address a defect of the original signal model despite its heuristic derivation.

The modification is retained throughout the rest of the work described in this thesis, but will not generally feature in the mathematics for reasons of notational clarity. It should, however, be borne in mind that the same principle as described in (i) to (iii) in section 4.2 is used everywhere that data is passed up or down the quadtree, be it gray level data or any other kind.

Those figures which show all levels of the quadtree will henceforth show two images for each of levels 1 to $Y-1$, the additional image being that formed at the interstitial nodes on each level. The images formed at the original nodes will appear in the left half of each figure, while those at the interstitial nodes will appear in the right half.

Note that there are no interstitial nodes on levels 0 (the root) or Y (the image plane).

With the removal of the blocking effect, it becomes rather easier to see the problem of the blurring of edges which remains. The estimator, in its spatially-invariant form, applies the same degree of smoothing everywhere in the image. Inevitably this implies excessive smoothing in the vicinity of nonstationarities such as edges and inadequate smoothing in regions of near-constant gray level. The former gives rise to blurring at nonstationarities while the latter allows excessive noise breakthrough in smoother regions. These complementary symptoms of the spatially-invariant processing are each visible — and objectionable to the viewer — in figures P11 and P12.

The following sections deal with the solution to this problem, which requires the introduction of spatially-variant processing.

4.4 Edge Preservation — Motivation

A problem with the estimator as it stands is that like most estimation schemes designed to remove noise, it tends to blur salient features such as lines and edges in the image. These features are known[9][54][55][56] to be of particular importance to the visual system, and an image in which the edges have been preserved will usually appear to the viewer to be of a higher quality than one in which the edges have been blurred. This may be true even when the blurred image has a superior signal-to-noise ratio (SNR).

A related visual phenomenon is the *masking effect* [13][105][158], which describes a reduction in the visibility of noise in the vicinity of rapid luminance changes such as edges.

A number of estimation schemes have been devised (see section 2.1) which attempt to take account of these effects. They seek to preserve edges by reducing the degree to which the image is smoothed in the vicinity of those edges. More noise therefore 'breaks through' the estimation near the edge, but the masking effect states that it will be less visible to the viewer precisely because of the sharpness of the preserved feature. (Anisotropic processing offers the prospect of increasing further the degree of noise smoothing without blurring edges — this will be treated in Chapters 5 and 6.)

However, edge detection plays an important role in image processing which is by no means confined to the adaptation of estimation strategies to visual characteristics as described above.

Many schemes for segmentation[51][130][131], feature extraction and 'image understanding' based on object-oriented descriptions[15][106], for example, use edge detection in some form, often as a 'front-end' or preprocessing system.

An effective edge detector, particularly one which is capable of some degree of noise rejection, is therefore a valuable tool in its own right, and one which could find wide applicability.

Many edge detectors[31][123][126] begin by calculating a luminance (gray level) gradient function across the image, and label as an edge any pixel at which this function is, for example, above a given threshold or at a local maximum.

Alternatively, they may calculate a second spatial derivative[90] and use its zero-crossings to define edge points. This method always gives edge maps in the form of closed contours[106].

More sophisticated systems[33] may estimate the local orientation of an edge (which is perpendicular to the direction of the local maximum luminance gradient) and, for example, attempt to link labelled nodes along the oriented direction and across gaps.

The bane of edge detectors, however, is noise[32][123]. Edges are rapidly-varying signals in the spatial domain and thus have significant components at high spatial frequencies. These properties unfortunately characterise noise rather well too, and a differential operator such as that used to extract the luminance gradient will respond enthusiastically to noise, often signalling false edges.

The edge detector may for this reason be preceded by (or, if it is linear, combined with) lowpass filters of one or more scales which smooth both the offending noise and the edge. As Canny[14] remarks, this provides a tradeoff between signal-to-noise ratio (SNR) at the output of the (linear) edge detector and its ability to localise the edge. Wilson and Granlund[146] observe that the same tradeoff is forced in segmentation, where it is manifested as a tradeoff between certainty of class membership and certainty in position. Both papers derive lower bounds on the product of the two competing parameters in the explicit form of the uncertainty principle[46].

Witkin[151] has considered the treatment of the scale at which edges are detected in images by proposing a 'scale-space' edge representation which makes explicit the relationship between components of an edge at different scales.

The use of edge detection operators of various scales is postulated by Marr[91] and Marr and Hildreth[90] as a model for edge detection in the visual system. They argue that if an edge (as indicated by a zero-crossing in the output of a second-spatial-derivative, or Laplacian, filter) is present over a range of scales (i.e. filter sizes) then it is probably genuine and not an artifact of the processing or a result of noise. This idea is taken up in section 4.5.2 below.

Eklundh et al [26] and Schunk[125] have adopted similar strategies in their use of multiple gaussian filters (Marr-Hildreth operators) of different bandwidths for edge detection.

All of the above-mentioned edge or gradient detectors utilise spatial operators which must be convolved with the noisy image data. This procedure involves a number of multiply-accumulate operations per image pixel which is equal to the size of the operator mask.

In the light of the computational savings achieved by using a quadtree-based vertical estimator, the question arises as to whether it is possible to design an adequate quadtree-based, vertical edge detector. In the following discussion, an heuristic argument is used to establish the general form of the system.

4.5 A Fast Vertical Edge Detector

Some heuristic requirements may be stated as follows:

(i) The method should possess a scale-invariant and recursive structure appropriate for implementation on the quadtree;

(ii) The subsequent modification of the estimator coefficients should preserve the range (from zero to unity) which those coefficients occupy;

(iii) The modified estimation scheme should preserve the basis of MMSE optimality from the spatially-invariant case.

4.5.1 Two Indices of Local Image Activity

(i) A Simple Index

At a node $(l+1, i', j')$ in the quadtree, an index of activity or 'busyness' may be defined as

$$\kappa_{l+1,i',j'} = \frac{(x_{l+1,i',j'} - x_{l,i,j})^2}{\hat{N}_{l+1}} \quad ,$$

$$(l+1, i', j') \in D_{l,i,j} \quad ,$$

$$0 \le l \le Y-1 \quad . \tag{4.4}$$

This index represents the energy of the link between a child and its father in the quadtree of noisy data if the link is considered to have a value equal to the difference between the values of the nodes. The index is normalised with respect to the average energy in all of the links between the two levels l and $l+1$ in question, which is given as $\hat{N}_{l+1}$ by (3.88), (3.91) and (3.92).

Thus a node which exhibits activity 'typical' of its level will have an activity index of unity. A node in a region of constant gray level will have an activity index close to zero, and a node which is in a 'busy' region such as an edge will have a high value of activity index.

(ii) A Propagated Index

A more complicated 'propagated' activity index may be defined. This index, at a given node, is a function of the 'simple' indices at nodes in the ancestor set of the given node and is defined recursively:

$$\alpha_{l+1,i',j'} = (\alpha_{l,i,j})^{(\theta_{l+1})}(\kappa_{l+1,i',j'})^{(1-\theta_{l+1})} ,$$

$$(l+1,i',j') \in D_{l,i,j} ,$$

$$0 \le l \le Y-1 , \qquad (4.5)$$

with initial condition

$$\alpha_{0,0,0} = 1$$

and where θ_{l+1} , $0 \le \theta_{l+1} \le 1$ is a weighting exponent which depends on the signal-to-noise ratio (SNR) of the noisy image (see below).

Taking logarithms in (4.5),

$$\log \alpha_{l+1,i',j'} = \theta_{l+1} \log \alpha_{l,i,j} + (1-\theta_{l+1}) \log \kappa_{l+1,i',j'} , \qquad (4.6)$$

a form which bears considerable resemblance to the linear estimator of (3.77).

A logarithmic domain for the measurement of image activity seems appropriate on the basis of the Weber-Fechner law[105] which observes that visual contrast (i.e. activity) sensitivity is a logarithmic type of measure, constant over many decades of illumination strength, as $\Delta I / I = K$ where ΔI is the detectable intensity change given a background intensity I, and K is a constant. The logarithmic domain has also proved useful in homomorphic signal processing[107][108]. Knutsson, Wilson and Granlund[75] use an

exponent of 0.5 in their 'linear feature' activity detector, which is not in practice dissimilar from the above logarithmic form. However, they employ only a single resolution of the detector.

4.5.2 Use of Weighted Geometric Mean for Noise Rejection

The propagated activity index $\alpha_{l+1,i,j}$ of (4.5) is a weighted geometric mean (WGM) of the 'simple' activity index $\kappa_{l+1,i,j}$ at the current node and the propagated activity index α_{l,i_o,j_o} at its father.

The WGM was chosen in preference to, for instance, an arithmetic (linear) combination on the basis (see section 4.5) that an edge persists over a number of contiguous levels of resolution (i.e. levels of the quadtree) whereas activity due to noise is reduced as the quadtree is ascended by the action of the averaging process. In order to achieve noise rejection, activity indices from different levels should therefore be combined in an 'AND' (geometric) rather than an 'OR' (arithmetic) fashion, and the WGM is indeed a weighted 'AND' combination.

The propagated activity index of (4.5) is affected by the combination exponent θ_k in such a way that the index is influenced more (less) by activity on upper levels of the tree if the SNR of the image data is lower (higher). Hence if the image is noisy, less account is taken of activity on the lower levels and more of activity on the upper levels in calculation of the propagated index.

The exponent θ_{l+1} estimates that proportion of the 'simple' activity index $\kappa_{l+1,i,j}$ which is due to noise, and decreases at higher levels (smaller l) as the noise is progressively averaged out.

$$\theta_{l+1} = \frac{3}{4}\frac{\sigma_v^2}{4^{Y-l-1}\hat{N}_{l+1}} \quad . \tag{4.7}$$

This value results from (3.83), which indicates that N_{l+1} contains a term $(3/4)(\sigma_v^2 / 4^{Y-l-1})$ which is due to the corrupting noise. (4.7) represents a normalisation of this term with respect to the sample estimate $\hat{N}_{l+1}$ of N_{l+1}.

The form of the propagation ensures that activity at a node will be represented in the propagated activity index to a greater extent if the ancestors of the node are also active, and to a lesser extent if they are not (the suspicion in this case being that the activity is due to noise).

For a noiseless image ($\sigma_v^2 = 0$), (4.7) gives $\theta_{l+1} = 0$, $0 \le l \le Y-1$. In this case there is no downward propagation and the propagated and 'simple' activity indices are equal at every node.

The exponentiation operation in (4.5) and the initial value of unity for $\alpha_{0,0,0}$, together with the normalisation in (4.4), ensure that a 'typical' node with an activity index $\kappa_{\ldots} = 1$, all the ancestors of which are also 'typical', will have a propagated activity index $\alpha_{\ldots} = 1$. The desirability of this will shortly become apparent.

Note that the activity index is a continuous variable. No attempt is made to assign binary labels ('edge'/'not edge') to any node since 'edginess' is intuitively a continuous attribute and is also a function of the scale of examination.

4.6 The Spatially-Variant Form of the Estimator

It now remains to determine how the propagated activity index $\alpha_{l,i,j}$ should affect the estimation. As noted in section 4.4, it is desirable to effect less smoothing of the image data in the vicinity of edges in order that those edges are not blurred. The penalty incurred is that more noise will remain in edge regions, but the masking effect indicates that its visibility will be reduced by the presence of the edge in close proximity.

Considering the recursive form of the estimator given in (3.77), it is apparent that in order to reduce the degree of smoothing it is necessary to reduce the estimator coefficient e_{l+1}, thus propagating less information $\hat{s}_{l,i,j}$ from above and retaining more of the noisy observation $x_{l+1,r,f}$.

Conversely, in smooth regions of the image where noise is at its most visible, it is desirable to increase the degree of smoothing and hence the estimator coefficient.

These objectives are achieved by replacing the spatially-invariant estimator coefficient e_{l+1} of (3.77) by a modified, spatially-variant coefficient

$$e_{l+1,r,f} = (e_{l+1})^{\alpha_{l+1,r,f}} \quad . \tag{4.8}$$

Neglecting the propagation of the activity index and considering the case where $\alpha_{\ldots,\ldots} = \kappa_{\ldots,\ldots} = (x_c - x_f)^2 / N$ where x_c is the child of the father x_f and N is the appropriate average of the link energies, there is a correspondence between the exponentiation of (4.8) and a gaussian form such as

$$e = e^{\frac{(x-\mu_x)^2}{\sigma^2}}$$

which may be considered as similar to the probability $P(x_c \notin S_E)$ of the node x_c not being in an edge region. However, this notion has been avoided in the development, since 'edginess' has been felt to be a continuous attribute as noted above.

(4.8) implies that for 'typical' nodes in the sense of section 4.5.1 which have $\alpha_{i+1,r,f} = 1$, the estimator coefficient is unchanged, while at more active nodes for which $\alpha_{i+1,r,f} > 1$ it is reduced toward zero and at less active nodes for which $\alpha_{i+1,r,f} < 1$ it is increased toward unity.

This illustrates the desirability of the forms of (4.4), (4.5) and (4.8) in that:

(i) a 'typical' node is still estimated using the original optimal MMSE estimator coefficient e_{i+1} and so the modification preserves the existing basis of MMSE optimality from the spatially-invariant case;

(ii) The (hypothetical) extreme values of the activity index (i.e. 0 and ∞) map on to appropriate extreme values of the modified estimator coefficient (i.e. 1 (total smoothing) and 0 (no smoothing)) with no risk of possible values of the activity index causing inadmissible values of the estimator coefficient (i.e. $\varepsilon_{i+1,r,f} < 0$ or $\varepsilon_{i+1,r,f} > 1$). Such values are inadmissible because, as is apparent from equation (3.79), this would imply negative values of β_k^2. Thus the modification of the estimator coefficient is required to be a mapping $[0,1] \rightarrow [0,1]$ of the closed unit interval; this is achieved by (4.8).

There is a marked correspondence with the 'masking function' and the 'visibility function' introduced by Anderson and Netravali[6] and used by Abramatic and Silverman[3] as the means of introducing spatial adaptivity into their filtering scheme. The present

activity index is analogous to their masking function, and their visibility function was used in much the same way as the estimator coefficient here in that it controlled the proportions of filtered and unfiltered data in the final estimate. They achieved the mapping of the masking function (MF) on to the visibility function (VF), confined to the unit interval [0,1], by using a mapping of the form

$$VF = \frac{1}{1+\gamma(MF)}$$

which is very similar to the present mapping of the activity index on to the estimator coefficient.

Appendix 4 contains a conference paper[19] which presents the nonstationary estimator; the derivation of the stationary method is similar to that of Appendix 1. The paper represents an early stage of development, however, and many of the indices have subsequently been altered.

4.7 The 'Signal-Equivalent' Spatially-Variant Model

Equations (3.78) and (4.8) are reproduced for reference:

$$e_{l+1} = \frac{\frac{\beta_{l+2}^2}{4} + \cdots + \frac{\beta_\gamma^2}{4^{\gamma-l-1}} + \frac{\sigma_v^2}{4^{\gamma-l-1}}}{\beta_{l+1}^2 + \frac{\beta_{l+2}^2}{4} + \cdots + \frac{\beta_\gamma^2}{4^{\gamma-l-1}} + \frac{\sigma_v^2}{4^{\gamma-l-1}}} , \qquad (3.78)$$

$$e_{l+1,r,f} = (e_{l+1})^{a_{l+1,r,f}} . \qquad (4.8)$$

Following (3.78), it is possible to introduce a 'signal-equivalent' characterisation of the

edge information (as expressed in the form of the modified estimator coefficient $\varepsilon_{l+1,\ell,f}$ of (4.8)) which is analogous to that of Abramatic and Silverman[3] and Knutsson, Wilson and Granlund[75]. This involves the replacement of the spatially-invariant model parameters β_k^2 with spatially-variant parameters $\beta'^2_{k,m,n}$ such that (suppressing position indices on $\beta'_{\ldots}$ for clarity)

$$\varepsilon_{l+1,\ell,f} = (e_{l+1})^{\alpha_{l+1,\ell,f}} = \frac{\dfrac{\beta'^2_{l+2}}{4} + \cdots + \dfrac{\beta'^2_Y}{4^{Y-l-1}} + \dfrac{\sigma_v^2}{4^{Y-l-1}}}{\beta'^2_{l+1} + \dfrac{\beta'^2_{l+2}}{4} + \cdots + \dfrac{\beta'^2_Y}{4^{Y-l-1}} + \dfrac{\sigma_v^2}{4^{Y-l-1}}} \tag{4.9}$$

which is the solution to the MMSE estimation problem for the spatially-variant model of (2.6) with $\beta_{l+1,\ell,f}$ replaced by $\beta'_{l+1,\ell,f}$.

Note that for $\varepsilon_{Y,m,n}$, (4.9) gives

$$\varepsilon_{Y,m,n} = \frac{\sigma_v^2}{\beta'^2_{Y,m,n} + \sigma_v^2} = (e_Y)^{\alpha_{Y,m,n}} \tag{4.10}$$

and from (3.79),

$$e_Y = \frac{\sigma_\zeta^2}{\beta_Y^2 + \sigma_v^2} \ . \tag{4.11}$$

Eliminating e_Y gives $\beta'^2_{Y,m,n}$ as

$$\beta'^2_{Y,m,n} = \sigma_v^2 \left[\left[1 + \frac{\beta_Y^2}{\sigma_v^2} \right]^{\alpha_{Y,m,n}} - 1 \right] \ . \tag{4.12}$$

In edge regions where $\varepsilon_{Y,m,n} < e_Y$, it follows from (4.10) and (4.11) that

$\beta'^2_{\gamma,m,n} > \beta^2_\gamma$, and so in edge regions the spatially-variant model coefficients are larger than the spatially-invariant coefficients.

Conversely, in smooth image regions where $\varepsilon_{\gamma,m,n} > e_\gamma$, (4.10) and (4.11) yield $\beta'^2_{\gamma,m,n} < \beta^2_\gamma$, and the values of the spatially-variant model coefficients are smaller than in the spatially-invariant case.

For 'typical' nodes (γ, m, n) with activity index $\alpha_{\gamma,m,n} = 1$, (4.10) gives $\varepsilon_{\gamma,m,n} = e_\gamma$ and (4.12) gives $\beta'^2_{\gamma,m,n} = \beta^2_\gamma$, unchanged from the spatially-invariant case.

With the $\beta'^2_{\gamma,\ldots}$ determined as in (4.12), it is possible to calculate the $\beta'^2_{\gamma-1,\ldots}$ from $\varepsilon_{\gamma-1,\ldots}$ and $e_{\gamma-1}$ and so on, giving a backward recursion for the spatially-variant model parameters $\beta'_{\gamma,\ldots}$.

4.8 Edge Detector Examples

Examples of the operation of the edge detector are shown in figures P13 to P15 (see Table 4.2).

Figure P13 shows the 'simple' (unpropagated) activity index $\kappa_{\gamma,\ldots}$ as given by (4.4) for the 'blobs' image at an input signal-to-noise ratio of 0dB. In the display, black corresponds to zero and white corresponds to the maximum value attained on the entire quadtree.

Figure P14 shows the propagated activity index $\alpha_{\gamma,\ldots}$ of (4.5) for the same source image and display conditions. The reduction in the level of spurious responses due to noise as compared to the 'simple' index is evident.

Figure P15 shows the estimator coefficient as modified according to (4.8) on a scale where black corresponds to zero and white to unity. The reduction of the estimator coefficient in edge regions is apparent.

Simple index, 0dB	P13
Propagated index, 0dB	P14
Estimator coefficient, 0dB	P15

Table 4.2

Edge Detector Examples

The edge detector evidently trades off resolution (equivalent to the measures of edge localisation used by Canny[14] or Wilson and Granlund[146]) against noise. At poorer input SNRs, the weighted geometric mean of (4.5) is constructed in such a way that the combination favours activity detected at larger scales, which is apparent for example in the spreading of the edges on the lower quadtree levels (i.e. toward the image plane) in figure P14.

The weighted geometric mean has an effect upon the propagated activity index which is rather similar to that of the estimator proper on gray level data as regards the increased smoothing (corresponding to poorer localisation of edges) which is applied at poorer signal-to-noise ratios. The performance of the system in the presence of noise is thus limited by a form of uncertainty relation as discussed in section 4.4.

4.9 Advantages of the Edge Detector

The edge detector described above inherits many of the advantages which accrue to the estimator as a result of its definition on the quadtree structure. It operates vertically and does not require the use of a spatial operator as do most edge detection schemes extant[31][90][123][126].

The computational burden amounts to three multiplications, one division and one exponentiation per node for calculation of the propagated activity index. Modification of the estimator coefficient demands a further exponentiation at each node.

If calculation of the activity indices is performed in the log domain, the requisite computations are one logarithm, 5/4 multiplications on average, and one exponentiation per node. Again, the modification of the estimator coefficient requires an additional exponentiation.

Calculation of the 'simple' activity index $\kappa_{....}$ may be performed in parallel at all nodes in the quadtree simultaneously. Calculation of the propagated index $\alpha_{....}$ may be performed in parallel at all nodes on a particular level, there being a level-to-level 'causality' analogous to that of the gray-level estimator due to the downward propagation of information.

The estimator coefficients may all be modified simultaneously.

Thus the edge detector, like the estimator, is highly amenable to parallel computation.

The notions of scale invariance and scale independence from the image model and esti-

mator also apply to the edge detector, which detects 'candidate' edges at different scales isomorphically and then combines the information to achieve noise rejection.

The structural similarity of the edge detector and image model/estimator results in their being easily combined to yield the spatially-variant ('signal-equivalent') form of the model, for which the modified estimator is the optimal MMSE solution.

Calculation of the edge information and the gray-level estimate at a given node may be combined into one conceptual 'operation' which is identical in structure at every node.

4.10 Estimation Results for the Full Implementation

The results referred to in this section were obtained using the estimator with both of the improvements (interstitial nodes and activity detection controlling spatially-variant estimation) described in this chapter.

4.10.1 Examples of Estimated Images

Figures P7, P8 and P16 — P31 show noisy input and estimated output images. Table 4.3 provides a key to the results presented. Only the image plane of the estimated quadtree is shown, except in the case of the 'girl' image at input SNR of 0dB, where the whole quadtree is shown in figure P18.

Image	'Girl'	'Blobs'	'Lake'
Noisy, 12dB	P16	P21	P26
Estimated, 12dB	P17	P22	P27
Noisy, 0dB	P7	P8	P28
Estimated, 0dB	P18	P23	P29
Noisy, -12dB	P19	P24	P30
Estimated, -12dB	P20	P25	P31

Table 4.3

Noisy and Estimated Images

4.10.2 Improvement of Signal-to-Noise Ratio

Table 4.4 shows the signal-to-noise ratio of the estimator output against that of the input for the various test images.

Image	'Girl'	'Blobs'	'Lake'
Input SNR		Output SNR	
12dB	18.8dB	22.1dB	17.2dB
0dB	13.6dB	15.0dB	12.0dB
-12dB	8.5dB	9.5dB	7.2dB

Table 4.4

Input and Output Signal-to-Noise Ratios

Comparison of the output SNRs for the restored 'girl' and 'blobs' images at 0dB input SNR with those attained by the spatially-invariant estimator of section 4.2 shows that further numerical improvements of 1dB and 2dB in the 'girl' and 'blobs' images respectively have resulted from the use of the spatially-variant scheme.

The improvement of 15dB at 0dB input SNR which is achieved by the present method for the 'blobs' image compares very favourably with the 8.3dB achieved by the reduced-update Kalman filter (RUKF) of Woods[156][157] and with the 9.3dB achieved by the semicausal filtering model of Jain[61]. The 'doubly-stochastic gaussian' (DSG) model of Woods, Dravida and Mediavilla[152] using the 'girl' image at an input SNR of 12dB attained an improvement of 4.6dB for the RUKF and 5.8dB for their 'M-algorithm'. The present scheme achieves an improvement of 6.8dB for this image at 12dB input SNR. It is implied in [152] that the DSG and RUKF methods may begin to break down at around 3dB input SNR; this is not true of the present system, which functions respectably at very low signal-to-noise ratios.

4.10.3 Discussion of Results

The effect of the introduction of spatially-variant processing may be seen by comparing the spatially-invariant estimates of the 'girl' and 'blobs' images in figures P11 and P12 respectively with their spatially-variant counterparts, figures P18 and P23; the changes are probably less confused by other detail in the case of the 'blobs' image.

The reduction in the level of noise in the smooth regions by comparison with the spatially-invariant scheme is immediately apparent and is a consequence of the low (less than unity) activity index in these areas giving rise to an increased estimator coefficient and hence to a greater degree of smoothing. In the vicinity of the object boundaries,

however, there is clearly less smoothing than in the spatially-invariant case. This is a result of high values of the activity index causing a reduction in the estimator coefficient. The sharpness of the boundaries is greatly enhanced.

The breakthrough of noise in the boundary regions is visible; close inspection of these areas reveals that small-scale, high-frequency noise is prevalent at the boundary while somewhat further (say about 1mm) away the noise is of larger scale or lower frequency. This is due to the scale-invariant operation of the estimator, noise processes at larger scales being propagated down from the boundary region at higher levels of the quadtree.

All of these effects of the spatially-variant scheme are visible also in the 'girl' image of figure P18 by contrast with figure P11. There is clearly less noise in the smooth regions and more in the edge regions, with a marked sharpening of the edges.

It should be mentioned that there is a constant level of noise across the spatially-invariant estimates of figures P11 and P12. Any apparent variation is due to the characteristics of the display equipment, photographic film and/or the visual properties of the viewer.

Comparison of the estimated 'girl' images of figures P17, P18 and P20 or of the estimated 'blobs' images of figures P22, P23 and P25 show the effect of variation of the input signal-to-noise ratio (SNR). At fairly high input SNR (12dB) the residual noise is of small scale and is confined to a narrow band centered on the edges in the image. As the input SNR is reduced, the affected band widens and the residual noise becomes of larger scale, as noted above, toward the band edges. This effect results from the widening of the edge bands, as indicated by the propagated activity index, as the input image is progressively degraded. This is a consequence of the increased uncertainty in edge position as the noise level increases. In fact, the residual noise even exactly at the edge posi-

tion becomes of lower frequency or larger scale as the input SNR is reduced because the estimator coefficient is affected less by activity at the lower levels for lower SNR. At very low input SNR (-12dB, figures P20 and P25) the residual noise is almost exclusively of large scale — hence its somewhat blotchy appearance — and the noise bands around the edges are wide enough to cause noise breakthrough over much of the image.

The degradation is graceful, however, and the estimates displayed as figures P18, P25 and P31 at the distinctly poor input SNR of -12dB are rather acceptable by comparison with what might be expected from existing estimation schemes. Unfortunately there are no published results, as far as the author is aware, for other schemes at such low SNR.

The results at input SNRs of 0dB and 12dB may, however, be compared with published work.

The present scheme gives results which are numerically and subjectively considerably superior to those achieved by the reduced-update Kalman filter of Woods[156][157]. The output of this latter method at an input SNR of 0dB exhibits a large amount of blurring and streaking and is of poor subjective quality. The semicausal filtering model of Jain[61] gives a result at 12dB input SNR which is distinctly blurred as a result of the stationarity of the estimator, although the noise level is reduced. This demonstrates well the need for spatially-variant processing.

The calculation of the activity index in the current scheme is such that at lower input SNR, less account is taken of activity on the lower quadtree levels. This results in more smoothing in the estimation of the lower levels near or at an edge, and so fine detail tends to be lost at low input SNR; this effect is in addition to the greater degree of smoothing applied generally at low SNR because of the higher value of the stationary estimator

coefficient.

Perhaps rather surprisingly, the leaves of the trees in the foreground of the restored 'lake' image are little affected by the change in input SNR from 12dB to 0dB. This is because the whole tree region generates a moderately high value of activity index at the middle levels of the quadtree in both cases, resulting in little smoothing at the scale of the leaves but more at the scale of the noise on the lower levels.

The design of the scheme is such that noise breakthrough occurs in the vicinity of edges. As discussed in sections 1.5 and 2.1, edges are locally one-dimensional and have orientation as a defining attribute. Following the successful application of anisotropic filtering based on local orientation to image restoration by Knutsson, Wilson and Granlund[75], it is reasonable to suppose that it may be possible to improve the restorations by utilising anisotropic filtering in the edge regions. The extension of the estimator to encompass anisotropic processing is the subject of Chapter 6.

It should be emphasised that the modifications described in this chapter are motivated and derived heuristically, and that the modified estimator is not optimal in the MMSE sense for the given signal model. However, the modifications clearly achieve a distinct improvement in both the visual and the numerical quality of the results, and this suggests that they represent an improvement over the original signal model and MMSE error criterion although the modifications are not expressed formally as a signal model and error criterion for which the optimal estimator might be designed. A disadvantage with that optimal approach is that the simple recursive structure would in all likelihood be lost; the modifications presented in this chapter preserve the computationally efficient form of the scheme.

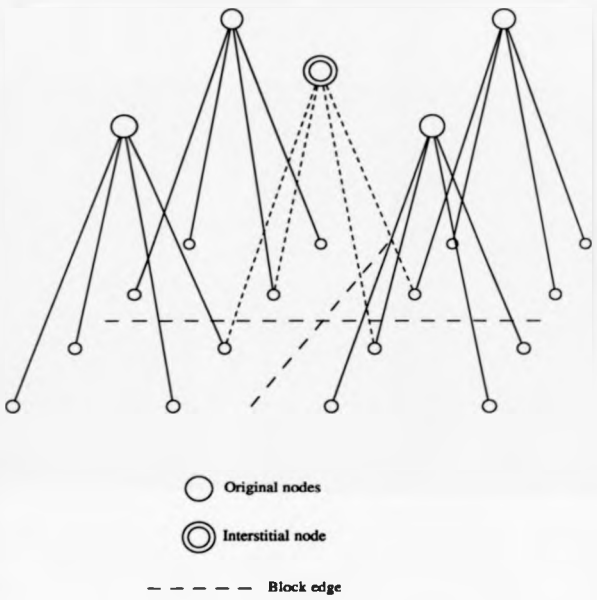


Figure 4.1 — Interstitial Node Positioning

CHAPTER 5

VECTOR ESTIMATION AND ORIENTATION

5.1 Motivation

A quadtree-based image model and its associated optimal MMSE estimator were presented in Chapters 2 and 3 respectively. Chapter 4 additionally described a modification to the estimator which removes the problem of block edge artifacts in the estimated image through the insertion of some additional nodes into the quadtree structure. Chapter 4 also introduced the spatially-variant form of the estimator, which was controlled by the output of a fast edge (or 'activity') detector operating on the quadtree.

The development has been based thus far on the assumption of scalar data. However, in certain applications the data to be processed may take the form of a vector-valued, rather than a scalar (gray level), field. Such vector fields might arise, for example, in the parametric description of textures[44][131], as a representation of the orientation of image features[38][75][78], or more directly from colour or multispectral images.

A vector form of the estimator might find application in the restoration of such fields from corrupted observations, or in coding, where a vector form of the innovations estimate could be quantised for transmission or storage.

A vector form of the edge detector might be applied to texture segmentation problems where the texture field is described at each pixel by a feature vector[44][131] which is sufficiently complete as to represent those properties of the texture which are to serve as the basis for the segmentation. In this respect, such a scheme would be complementary to

that of Burt, Hong and Rosenfeld[12] (see section 2.2) who describe a segmentation strategy, applicable to vector data, which uses a pyramidal data structure and classifies on the basis of the parameters of region interiors whereas the presently postulated scheme would locate region boundaries.

Knutsson, Wilson and Granlund[75] have shown that anisotropic filtering, controlled by the estimated local orientation of image features such as lines and edges, gives improved visual results in image restoration because the filter can be orientation-tuned to be low-pass in the direction parallel to the feature and to be all-pass in the perpendicular direction. The 'sharpness' of the feature, which percept may be regarded as conveyed by oriented highpass components in the signal, is preserved while noise is smoothed parallel to it.

The authors of [75] further demonstrated (see their figure 2) that such parallel-filtered noise in the neighbourhood of a feature may enhance the visual detectability of the feature while noise filtered in the perpendicular direction will reduce it.

In order effectively to apply such techniques it is necessary to estimate the local orientation of the image from the noisy data, and so the orientation estimate (which is a vector field) will itself be noisy. Thus it might benefit from vector restoration before it is used in the anisotropic estimation of the scalar or vector image.

As will be shown in the following sections, the present scheme may be extended to vector data. The extended method with certain modifications may be used in the anisotropic estimation task.

5.2 Extension to Vector Data

5.2.1 Vector Form of the Model

The model of (2.6) generalises in an obvious way to the case where the nodes of the quadtree are vector-valued:

$$\mathbf{s}_{l+1,i',j'} = \mathbf{s}_{l,i,j} + \beta_{l+1,i',j'}\mathbf{w}_{l+1,i',j'} \quad ,$$

$$(l+1,i',j') \in D_{l,i,j} \quad ,$$

$$0 \le l \le Y-1 \quad , \tag{5.1}$$

$$\mathbf{s}_{0,0,0} = \beta_{0,0,0}\mathbf{w}_{0,0,0} \quad , \tag{5.2}$$

$$E\,\mathbf{w}_{l,i,j}\mathbf{w}_{r,m,n}^{T} = \mathbf{I}\,\delta_{l,r}\delta_{i,m}\delta_{j,n} \tag{5.3}$$

where $\mathbf{v}^T$ denotes the transpose of $\mathbf{v}$ and $\mathbf{I}$ is the identity matrix of appropriate order.

These equations define a vector martingale process where the innovations vector process $\mathbf{w}_{l,i,j}$ has components which are mutually orthonormal and are orthonormal to all components of all other $\mathbf{w}_{r,m,n}$.

If $\mathbf{s}_{\ldots}$ and $\mathbf{w}_{\ldots}$ are M-dimensional vectors, this model may be decomposed trivially into M concurrent scalar models of the form of (2.6).

If the correlation matrix of the vector innovations process were given, instead of by (5.3), by

$$E\,\mathbf{w}_{l,i,j}\mathbf{w}^{T}_{r,m,n} = \mathbf{Q}_w \delta_{l,r}\delta_{i,m}\delta_{j,n} \tag{5.4}$$

where

$$\mathbf{Q}_w = \mathrm{diag}(\lambda_k) \ ,$$

$$0 \le k \le M-1 \ , \tag{5.5}$$

then the model (5.1) decomposes into M scalar models of the form

$$s_{l+1,i',j'}(k) = s_{l,i,j}(k) + \sqrt{\lambda_k}\,\beta_{l+1,i',j'}\,w_{l+1,i',j'}(k) \ ,$$

$$(l+1,i',j') \in D_{l,i,j} \ ,$$

$$0 \le l \le \gamma-1 \ ,$$

$$0 \le k \le M-1 \ , \tag{5.6}$$

which is equivalent to the model of (2.6) with $\beta_{l+1,i',j'}$ replaced by $\sqrt{\lambda_k}\,\beta_{l+1,i',j'}$. The index k denotes the model for the k^{th} vector component.

Finally, if the correlation matrix of the vector innovations process were given by

$$E\,\mathbf{w}_{l,i,j}\mathbf{w}^{T}_{r,m,n} = \mathbf{Q}_w \delta_{l,r}\delta_{i,m}\delta_{j,n} \tag{5.7}$$

with $\mathbf{Q}_w$ a general positive definite correlation matrix, then the model decomposes into M scalar models of the form of (5.6) with $\sqrt{\lambda_k}$ replaced by $\sqrt{(\mathbf{Q}_w)_{k,k}}$.

Attention will be restricted at present to the spatially invariant form of the model of (5.1) — (5.3), with $\beta_{l+1,i',j'} = \beta_{l+1}$ in (5.1). The extension to the spatially-variant form will be treated in section 5.2.3.

5.2.2 Vector Form of the Estimator

For the model of (5.1) — (5.3), the linear MMSE estimator is a trivial extension of the scalar estimator of (3.77):

$$\hat{\mathbf{s}}_{l+1,i',j'} = e_{l+1}\hat{\mathbf{s}}_{l,i,j} + (1-e_{l+1})\,\mathbf{x}_{l+1,i',j'} \quad ,$$

$$(l+1,i',j') \in D_{l,i,j} \quad , \tag{5.8}$$

with e_{l+1} given again by (3.78).

The average-value data quadtree $\mathbf{x}_{\cdot\cdot\cdot\cdot}$ is formed exactly as in the scalar case, each node being set equal to the vector average of its four children.

It has been assumed here that each vector component is corrupted at level Y of the quadtree by additive white gaussian noise (AWGN) of variance σ_v^2 as previously. Were the corruption to vary between components, a different estimator (of the same form) would be required for each component. It might, however, be preferable even in this case to use the same estimator of (5.8) and to tolerate the resulting suboptimality with respect to each vector component, since in this way the balance between the various components is preserved. Use of a different estimator for each component might result, for instance, in the appearance of colour band artifacts at luminance gradients if the vector field were a colour image.

Note that the use of interstitial nodes in the formation of the average-value and estimated data quadtrees as described in section 4.2 is extended trivially to the vector case.

5.2.3 Vector Form of the Edge Detector

In a similar manner it is possible to extend the edge or activity detector of section 4.5 to the case of vector data.

The 'simple' activity index $\kappa_{l+1,i',j'}$ of section 4.5.1 is now defined as the normalised energy of the vector difference between a node and its father in the average-value quad-tree,

$$\kappa_{l+1,i',j'} = \frac{|\mathbf{x}_{l+1,i',j'} - \mathbf{x}_{l,i,j}|^2}{\hat{N}_{l+1}} \quad ,$$

$$(l+1,i',j') \in D_{l,i,j} \quad , \tag{5.9}$$

where $\hat{N}_{l+1}$ is now the average over level $l+1$ of the numerator term.

The propagated activity index $\alpha_{l+1,i',j'}$ and its effect on the estimator coefficient e_{l+1} remain as given in sections 4.5.1 and 4.6 respectively.

The 'signal-equivalent' spatially variant form of the model of (5.1) — (5.3) is derived for the vector case as it was in section 4.7 for the scalar case, on the basis of the spatially-variant estimator coefficient

$$\varepsilon_{l+1,i',j'} = (e_{l+1})^{\alpha_{l+1,i',j'}} = \frac{\frac{\beta'^2_{l+2}}{4} + \cdots + \frac{\beta'^2_{\gamma}}{4^{\gamma-l-1}} + \frac{\sigma_v^2}{4^{\gamma-l-1}}}{\beta'^2_{l+1} + \frac{\beta'^2_{l+2}}{4} + \cdots + \frac{\beta'^2_{\gamma}}{4^{\gamma-l-1}} + \frac{\sigma_v^2}{4^{\gamma-l-1}}} \tag{5.10}$$

where

$$e_{l+1} = \frac{\frac{\beta^2_{l+2}}{4} + \cdots + \frac{\beta^2_{\gamma}}{4^{\gamma-l-1}} + \frac{\sigma^2_v}{4^{\gamma-l-1}}}{\beta^2_{l+1} + \frac{\beta^2_{l+2}}{4} + \cdots + \frac{\beta^2_{\gamma}}{4^{\gamma-l-1}} + \frac{\sigma^2_v}{4^{\gamma-l-1}}} \tag{5.11}$$

is the spatially-invariant estimator coefficient. Position indices have been suppressed from $\beta'_{\ldots,l}$ in (5.10) in the interests of clarity.

As mentioned in section 5.1, the vector estimator described above may be applied to the anisotropic estimation task. The scheme may be used directly for the restoration of noisy orientation data and may be modified to incorporate the anisotropic estimation itself, which is the subject of Chapter 6.

5.3 Orientation in Images and in Vision

It is an observable fact and (Marr[91] p.49) a consequence of the cohesiveness of matter that the visual environment contains dense material objects and that these objects tend over a range of scales to have distinct and continuous boundaries (see section 1.5). These boundaries map on to discontinuities in, for example, luminance or texture in images (including the retinal image).

There is a considerable body of physiological and psychophysical evidence which indicates that the early stages of mammalian visual systems contain mechanisms for the detection of such edges at particular orientations.

The physiological studies of Hubel and Wiesel[54][55] and Blakemore and Campbell[9], supported by the psychophysical work of Kulikowski and King-Smith[81][82] and Mostafavi and Sakrison[94], have shown that there are nerve cells (neurons) in area 17 of the

visual cortex which respond to linear features in the retinal image which have orientations lying within a narrow angular band. The detectors form a quasi-regular array, indexed by orientation, which is replicated for different scales and positions within the visual field. Daugman[23][24] has examined the frequency-domain behaviour of such detectors and has proposed some formal properties by consideration of their spectral consequences. Marcelja[89] has explored mathematical descriptions of the detector responses.

As mentioned in section 1.5, Linsker[85] has shown that an artificial perceptual or 'neural' network of appropriate design, and operating under a simple learning rule, may develop a similar array of orientation-selective feature detectors when presented with a set of input patterns or even random noise fields.

The proposition that edges and their orientations are of considerable utility in image decomposition is thus supported by much evidence from real biological vision systems and even from relatively crude artificial simulations. The subjective percept of poor image quality resulting from the blurring of edges (a percept familiar, for example, to regular viewers of slide and overhead projections) provides further evidence for the importance of oriented, locally-linear features in visual perception.

5.4 Extraction of Orientation Information

This section is based largely on the work of Hans Knutsson[75-80].

The objective is the design of a filtering scheme which estimates the local orientation (i.e. within the filter support) of oriented image features.

A minimal but complete representation of orientation is a two-dimensional vector which indicates the orientation by its argument and the amplitude of the oriented feature by its modulus.

This is necessary to avoid the problem that the limit of any oriented feature as its magnitude (contrast) is reduced to zero is a constant gray level, and so a simple argument-only representation could take on any arbitrary value of orientation in a flat or isotropic region.

The vector representation, however, is continuous at this 'singularity' and the magnitude is of potential utility in controlling the degree of anisotropic filtering subsequently applied.

The desired filter, therefore, produces this vector field as its output.

5.4.1 Specifications for the Filter Design

The filter should possess the following characteristics:

(i) The magnitude of the vector response to an isotropic input should be zero;

(ii) The magnitude of the vector response for a given input energy should be maximal when the input function is 'maximally anisotropic', i.e. of one-dimensional variation on the filter support;

(iii) The magnitude of the vector response should be invariant with respect to the position of the input feature within the filter support. More generally, any input function

of one-dimensional variation and of given energy should produce the same magnitude of response.

(iv) The argument of the vector response should be unbiased in the sense that it should exactly equal the input orientation for input signals of one-dimensional variation on the filter support.

(v) The filter should be of smooth variation in the spatial frequency domain in order to avoid ringing and large significant extent in the spatial domain. (The filter must be spatially localised if the notion that its output represents *local* orientation is to be at all useful).

Property (iv) above implies that interpolation over the outputs of a number of filters, each maximally sensitive to a different orientation of the input feature, will be required. It will be shown in the development which follows that four is a suitable number of filters.

5.4.2 Use of Quadrature Filters

Property (iii) above implies that the filters should be insensitive to the phase (i.e. position) of the input feature and should have an all-pass type of transfer function in the frequency domain (so that any input function of one-dimensional variation and of given energy and orientation will produce the same output energy regardless of its composition in the spatial or equivalently in the spatial frequency domain).

The phase-insensitivity requirement means that the filters must each be a combination of an even and an odd filter in the spatial domain. The spatial impulse response of the even filter may be expressed as a sum of cosine functions,

$$h_e(\mathbf{y}) = \frac{1}{\pi} \int_0^{\infty} H_e(\boldsymbol{\omega}) \cos \boldsymbol{\omega}^T \mathbf{y} \, d\boldsymbol{\omega} \tag{5.12}$$

where

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

and

$$\boldsymbol{\omega} = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}$$

are spatial- and frequency-domain vectors respectively.

Phase insensitivity demands that the odd filter contain the same 'amount' of $\sin \boldsymbol{\omega}^T \mathbf{y}$ as does the even filter of $\cos \boldsymbol{\omega}^T \mathbf{y}$ for each $\boldsymbol{\omega}$, so that a sinusoidal input function of the form $\cos(\boldsymbol{\omega}^T \mathbf{y} + \phi)$ excites the odd filter as much for $\phi = (2n+1)\frac{\pi}{2}$ as it does the even filter for $\phi = 2n\frac{\pi}{2}$.

Thus the odd filter must be given by

$$h_o(\mathbf{y}) = \frac{1}{\pi} \int_0^{\infty} H_e(\boldsymbol{\omega}) \sin \boldsymbol{\omega}^T \mathbf{y} \, d\boldsymbol{\omega} \; . \tag{5.13}$$

A filter pair such as that described above is known as a *quadrature pair* since each member may be derived from the other by a Hilbert transform, which effects a phase shift of every frequency component by $\frac{\pi}{2}$ radians.

Quadrature filters have been used extensively in applications where there is a requirement of phase insensitivity. They appear in communication theory, for example as components of single-sideband (SSB) modulation schemes[137].

The studies of Movshon, Thompson and Tolhurst[95][96], Enroth-Cugell and Robson[30], and Pollen and Ronner[115][116] have identified 'complex' cells in the visual cortex which, while responding maximally to an oriented stimulus at the retina within their receptive fields like their 'simple' counterparts, are insensitive to its position within the field. This fact, combined with the known properties of the 'simple' cortical cells which are position-sensitive, are of both odd ('sine') and even ('cosine') receptive field varieties, and are known to feed information to the 'complex' cells, constitutes *prima facie* evidence for the existence of something approximating quadrature filter pairs in the early stages of vision.

Pollen & Ronner[115][116] in fact discovered in their experimental data twelve pairs of adjacent simple cells, each pair possessing the same orientation (within five degrees) and spatial frequency (within 1/4 octave) tuning, with the intra-pair spatial phase displacement having a mean value of 92.1 degrees and standard deviation of 6.5 degrees. This is perhaps the strongest evidence yet for the quadrature pair hypothesis in visual orientation detection.

5.4.3 Quadrature Filter Design

Rewriting (5.12) and (5.13) with infinite integration limits,

$$h_e(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H_e(\omega) \cos \omega^T y \, d\omega \qquad (5.14)$$

and

$$h_o(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H_o(\omega)\, \mathrm{sgn}(\omega) \sin \omega^T y \, d\omega \tag{5.15}$$

where $\mathrm{sgn}(\omega)$ is the binary-valued sign function. Its two-dimensional form will be considered shortly.

Expressing a Fourier transform pair as $h(y) \leftrightarrow H(\omega)$, it follows from (5.14) and (5.15) that for the even filter,

$$h_e(y) \leftrightarrow H_e(\omega) \tag{5.16}$$

and for the odd filter,

$$h_o(y) \leftrightarrow -j\, \mathrm{sgn}(\omega) H_e(\omega) \quad . \tag{5.17}$$

The various symmetries in the spatial and frequency domains may now be made explicit. For the even filter,

$$H_e(\omega) = H_e(-\omega) = \mathrm{Re}\, H_e(\omega) \quad , \tag{5.18}$$

$$\mathrm{Im}\, H_e(\omega) = 0 \quad , \tag{5.19}$$

$$h_e(y) = h_e(-y) = \mathrm{Re}\, h_e(y) \quad , \tag{5.20}$$

and

$$\mathrm{Im}\, h_e(y) = 0 \quad . \tag{5.21}$$

For the odd filter,

$$H_o(\omega) = -j\,\mathrm{sgn}(\omega) H_e(\omega) \tag{5.22}$$

$$H_o(\omega) = -H_o(-\omega) \ , \tag{5.23}$$

$$\mathrm{Re}\, H_o(\omega) = 0 \ , \tag{5.24}$$

$$h_o(\mathbf{y}) = -h_o(-\mathbf{y}) \ , \tag{5.25}$$

and

$$\mathrm{Im}\, h_o(\mathbf{y}) = 0 \ . \tag{5.26}$$

Returning to the definition of the function $\mathrm{sgn}(\omega)$, it is apparent from (5.22) that in two dimensions this function defines a reference axis for the filter although there seems no obvious candidate for the bisector of the ω–plane into 'positive' and 'negative' half-planes. It is, however, precisely this parameter which distinguishes filters tuned to different orientations, and so it will be defined *differently* for each orientation of filter. This need not cause confusion since the terms which include this function will be written explicitly in terms of a particular direction.

If the vector ω is represented as

$$\omega = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = |\omega| \exp(j\theta) \tag{5.27}$$

where

$$\theta = \tan^{-1}\frac{\omega_2}{\omega_1} \tag{5.28}$$

then sgn(ω) may be expressed as

$$\mathrm{sgn}(\omega) = \mathrm{sgn}(\cos\theta) \tag{5.29}$$

or more generally for arbitrary θ_k ,

$$\mathrm{sgn}(\omega) = \mathrm{sgn}(\cos(\theta - \theta_k)) \tag{5.30}$$

where the sign function has its more familiar one-dimensional form

$$\mathrm{sgn}(t) = \begin{cases} -1 \ , & t < 0 \ , \\ 0 \ , & t = 0 \ , \\ 1 \ , & t > 0 \ . \end{cases} \tag{5.31}$$

The effect of (5.30) is to partition the ω–plane into half-planes in which ω is nominally positive and nominally negative respectively. The boundary is the ω_2 axis, rotated through θ_k in the anticlockwise direction (see figure 5.1).

From the all-pass requirement of (iii), section 5.4.1, it follows that $H_a(\omega)$ should equal unity for all ω. However, as will become clear this violates the requirement of (v), section 5.4.1, that the filter should be of smooth frequency-domain variation and so the term $H_a(\omega)$ will remain until after consideration of the angular response.

The design specifications of section 5.4.1 are expressed in polar (magnitude and argument) terms, the two parameters being treated separately. It is therefore appropriate to

design the filters to be polar separable, of the form

$$H(\rho,\theta) = H_\rho(\rho)H_\theta(\theta) \quad . \tag{5.32}$$

Daugman[23] has shown that anisotropic visual filters are likely also to possess such polar separability.

As observed above, in order to introduce the requisite orientation tuning, multiple filters are needed. The output vector will then be an interpolation over the outputs of the individual filters, each of which is tuned to a particular orientation.

Thus the k^{th} filter pair is given in the frequency domain by

$$H_{ak}(\omega) = H_\rho(\rho)H_{\theta_k}(\theta) \quad , \tag{5.33}$$

and

$$H_{ok}(\omega) = -j\,\mathrm{sgn}(\cos(\theta - \theta_k))H_\rho(\rho)H_{\theta_k}(\theta) \quad , \tag{5.34}$$

and their combination by

$$\begin{aligned} H_k(\omega) &= H_k(\rho,\theta) = H_{ak}(\omega) + jH_{ok}(\omega) \\ &= 2H_\rho(\rho)H_{\theta_k}(\theta)U(\cos(\theta - \theta_k)) \quad , \end{aligned} \tag{5.35}$$

where θ_k is the orientation of the k^{th} filter pair and $U(.)$ is the unit step function.

The frequency-domain response of the k^{th} filter pair to an input $F(\omega) = F(\rho,\theta)$ is then given by

$$G_{ek}(\rho,\theta) = H_{ek}(\rho,\theta)F(\rho,\theta)$$

$$= H_{\rho}(\rho)H_{\theta_k}(\theta)F(\rho,\theta) \tag{5.36}$$

and

$$G_{ok}(\rho,\theta) = H_{ok}(\rho,\theta)F(\rho,\theta)$$

$$= H_{\rho}(\rho)H_{\theta_k}(\theta)F(\rho,\theta)(-j\,\mathrm{sgn}(\cos(\theta-\theta_k))) \tag{5.37}$$

from which

$$G_k(\omega) = G_{ek}(\omega) + jG_{ok}(\omega)$$

$$= H_{ek}(\omega)F(\omega)(1+\mathrm{sgn}(\cos(\theta-\theta_k)))$$

$$= 2H_{ek}(\omega)F(\omega)U(\cos(\theta-\theta_k)) \tag{5.38}$$

Hence the quadrature filter pair transfer function $H_k(\omega)$ and the response $G_k(\omega)$ are *analytic signals* [111] which are nonzero only on the 'positive' half-plane given by

$$U(\cos(\theta-\theta_k)) > 0 \tag{5.39}$$

in the frequency domain.

From (5.38) it follows that the magnitude and phase of the response $G_k(\omega)$ of the k^{th} filter pair are given by

$$|G_k(\omega)| = 2H_{ek}(\omega)|F(\omega)|$$

$$= 2H_{\rho}(\rho)H_{\theta_k}(\theta)|F(\rho,\theta)| \tag{5.40}$$

and

$$\arg G_k(\omega) = \tan^{-1} \frac{H_{ak}(\omega) \operatorname{Im} F(\omega)}{H_{ak}(\omega) \operatorname{Re} F(\omega)}$$

$$= \arg F(\omega) \qquad (5.41)$$

which represents a generalised measure of signal phase in the region of support of the filter.

The relationship between the various signals described above is shown in one-dimensional form in figure 5.2.

It remains to establish the form of the angular frequency response. The following lemma concerns the frequency-domain behaviour of signals which are spatially of one-dimensional variation, as invoked extensively in the specifications of section 5.4.1.

Lemma

For signals of one-dimensional variation in the spatial domain, the frequency-domain energy is concentrated on a line through the origin at an orientation perpendicular to that of the spatial function.

Proof

Consider the two-dimensional function

$$f(x, y) = f(y\cos\phi - x\sin\phi) \ ,$$

$$0 \le \phi < \pi \ , \qquad (5.42)$$

which is of one-dimensional variation, its orientation being ϕ and the direction of its variation being perpendicular to ϕ as in figure 5.3. Then in the frequency domain,

$$F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-jux}e^{-jvy}\,dxdy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y\cos\phi - x\sin\phi)e^{-jux}e^{-jvy}\,dxdy. \tag{5.43}$$

Defining the variables s and t as

$$s = y\cos\phi - x\sin\phi \tag{5.44}$$

and

$$t = x\cos\phi + y\sin\phi \tag{5.45}$$

and substituting in (5.43) gives

$$F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s)e^{-ju(t\cos\phi - s\sin\phi)}e^{-jv(s\cos\phi + t\sin\phi)}dsdt$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s)e^{-jt(v\cos\phi - u\sin\phi)}e^{-js(u\cos\phi + v\sin\phi)}dsdt$$

$$= \int_{-\infty}^{\infty} f(s)e^{-jt(v\cos\phi - u\sin\phi)}ds \int_{-\infty}^{\infty} e^{-js(u\cos\phi + v\sin\phi)}dt \tag{5.46}$$

from which

$$F(u,v) = F(v\cos\phi - u\sin\phi)\,\delta(u\cos\phi + v\sin\phi) \tag{5.47}$$

and the proof is complete.

Considering now the exact interpolation requirement of (iv), section 5.4.1, and given a real input $f(x, y)$ of one-dimensional variation oriented at an angle of $\phi + \frac{\pi}{2}$, it is proposed that an angular filter function

$$H_{\theta_k}(\theta) = \cos^2(\theta - \theta_k) \tag{5.48}$$

is appropriate and allows exact interpolation of the orientation of $f(x, y)$.

From (5.47) it follows that the Fourier transform of $f(x, y)$ is given in polar coordinates by

$$F(\rho,\theta) = F_\rho(\rho)\,\delta(\theta-\phi) + F_\rho^*(\rho)\,\delta(\theta-\phi+\pi) \ ,$$

$$\rho > 0 \ ,$$

$$0 \le \theta < \pi \ ,$$

$$0 \le \phi < \pi \ , \tag{5.49}$$

where $F_\rho^*(\rho)$ is the complex conjugate of $F_\rho(\rho)$: $\mathrm{Re}\,F_\rho^*(\rho) = \mathrm{Re}\,F_\rho(\rho)$ and $\mathrm{Im}\,F_\rho^*(\rho) = -\mathrm{Im}\,F_\rho(\rho)$.

Now from (5.48) and (5.35),

$$H_k(\rho,\theta) = H_\rho(\rho)\cos^2(\theta-\theta_k)U(\cos(\theta-\theta_k)) \ ,$$

$$0 \le \theta < \pi \ , \tag{5.50}$$

and from (5.38) and (5.49),

$$G_k(\rho,\theta) = F(\rho,\theta)H_k(\rho,\theta)$$

$$= \begin{cases} F_\rho(\rho)H_\rho(\rho)\delta(\theta-\phi)\cos^2(\theta-\theta_k) \ , \ \cos(\phi-\theta_k) > 0 \\ F_\rho^*(\rho)H_\rho(\rho)\delta(\theta-\phi+\pi)\cos^2(\theta-\theta_k) \ , \ \cos(\phi-\theta_k) < 0 \end{cases} \tag{5.51}$$

or

$$G_k(\rho,\theta) = \begin{cases} G_\rho(\rho)\delta(\theta-\phi)\cos^2(\theta-\theta_k) \ , \ \cos(\phi-\theta_k) > 0 \\ G_\rho^*(\rho)\delta(\theta-\phi+\pi)\cos^2(\theta-\theta_k) \ , \ \cos(\phi-\theta_k) < 0 \end{cases} \tag{5.52}$$

Taking inverse transforms yields

$$g_k(x,y) = \begin{cases} g(x\cos\phi + y\sin\phi)\cos^2(\phi-\theta_k) \ , \ \cos(\phi-\theta_k) > 0 \\ g^*(x\cos\phi + y\sin\phi)\cos^2(\phi-\theta_k) \ , \ \cos(\phi-\theta_k) < 0 \end{cases}$$

or

$$g_k(x,y) = \begin{cases} g(x,y)\cos^2(\phi-\theta_k) \ , \ \cos(\phi-\theta_k) > 0 \\ g^*(x,y)\cos^2(\phi-\theta_k) \ , \ \cos(\phi-\theta_k) < 0 \end{cases} \tag{5.53}$$

Now if four filter pairs are used, with orientations θ_k given by

$$\theta_k = \theta_0 + \frac{k\pi}{4} \ ,$$

$$0 \le k \le 3 \ , \tag{5.54}$$

then, noting that $|g^*(x,y)| = |g(x,y)|$,

$$|g_0(x,y)| = |g(x,y)|\left[\frac{1}{2} + \frac{1}{2}\cos(2\phi - 2\theta_0)\right] \tag{5.55}$$

$$|g_1(x,y)| = |g(x,y)|\left[\frac{1}{2} + \frac{1}{2}\sin(2\phi - 2\theta_0)\right] \tag{5.56}$$

$$|g_2(x,y)| = |g(x,y)|\left[\frac{1}{2} - \frac{1}{2}\cos(2\phi - 2\theta_0)\right] \tag{5.57}$$

$$|g_3(x,y)| = |g(x,y)|\left[\frac{1}{2} - \frac{1}{2}\sin(2\phi - 2\theta_0)\right] \tag{5.58}$$

and

$$|g_2(x,y)| - |g_0(x,y)| = |g(x,y)|\cos(2(\phi + \tfrac{\pi}{2}) - 2\theta_0) \ , \tag{5.59}$$

$$|g_3(x,y)| - |g_1(x,y)| = |g(x,y)|\sin(2(\phi + \tfrac{\pi}{2}) - 2\theta_0) \tag{5.60}$$

Thus if

$$g_c(x,y) = |g_2(x,y)| - |g_0(x,y)| \tag{5.61}$$

and

$$g_s(x,y) = |g_3(x,y)| - |g_1(x,y)| \tag{5.62}$$

then

$$\mathbf{g}(x,y) = \begin{bmatrix} g_c(x,y) \\ g_s(x,y) \end{bmatrix} = |\mathbf{g}(x,y)| \begin{bmatrix} \cos(2(\phi+\frac{\pi}{2}) - 2\theta_0) \\ \sin(2(\phi+\frac{\pi}{2}) - 2\theta_0) \end{bmatrix} \tag{5.63}$$

and ϕ is exactly determined from $\mathbf{g}(x,y)$ as required by specification (iv), section 5.4.1. Recall that the spatial input function $f(x,y)$ was oriented at an angle of $\phi + \frac{\pi}{2}$, as expressed in $\mathbf{g}(x,y)$.

The vector $\mathbf{g}(x,y)$ is the desired output which describes the orientation of the image. The constant θ_0 (the offset angle for the orientation vector) is chosen as

$$\theta_0 = \frac{\pi}{8} \tag{5.64}$$

so that no bias is introduced into the estimate by the structure of the discrete lattice, which might cause a preference for vertical and horizontal directions were any of the filter pairs thus oriented.

The filter orientations are then as shown in figure 5.4.

Specification (v) of section 5.4.1 states that the filters should be of smooth frequency-domain variation. Such is obviously not true of the sign function of (5.30) and hence not true of the frequency response of the odd member of each quadrature filter pair as given by (5.34). There is in general a discontinuity at $\omega = 0$.

In order to overcome this problem, but bearing in mind the requirement that the filters be all-pass in character, the radial frequency response $H_\rho(\rho)$ is rolled off as a cosine-squared near the origin $\rho = 0$. The response reaches unity by $\rho = 0.05\pi$. A similar roll-off is applied between $\rho = 0.95\pi$ and $\rho = \pi$.

5.4.4 Double-Angle Vector Representation

Equation (5.63) shows that the orientation estimate is expressed in terms of $\cos 2\phi$ and $\sin 2\phi$ for a local image orientation of ϕ.

This parametric 'double-angle' representation[38][78] of ϕ overcomes the degeneracy of the inverse trigonometric functions which are one-to-many mappings. Restriction of ϕ to an interval $\phi_1 \leq \phi < \phi_1 + \pi$ (note that the orientation has period equal to π) creates an arbitrary discontinuity of magnitude π in the inverse trigonometric functions at $\phi = \phi_1$.

The functions $\cos 2\phi$ and $\sin 2\phi$, however, contain no such discontinuities and proceed through one cycle for each cycle of orientation.

5.5 Effect of Noise on the Orientation Estimate

The corruption of the image by additive white gaussian noise gives rise to a similar corruption in the frequency (DFT) domain.

Considering the one-dimensional case, if the corrupted data is given by

$$x(n) = s(n) + v(n) \tag{5.65}$$

where $s(n)$ is the signal and $v(n)$ is white noise,

$$E\,v(n)v(m) = \sigma_v^2 \delta_{n,m} \quad , \tag{5.66}$$

then in the DFT domain, by the linearity of the DFT,

$$X(k) = S(k) + V(k) \tag{5.67}$$

where

$$G(k) = N^{-0.5} \sum_{n=0}^{N-1} g(n) e^{-j2\pi \frac{nk}{N}} \tag{5.68}$$

is an energy-preserving DFT.

The correlation function of the noise component $V(k)$ of $X(k)$ is given by

$$EV(k)V^*(m) = N^{-1} \sum_{n=0}^{N-1} \sum_{l=0}^{N-1} e^{-j2\pi \frac{nk}{N}} e^{-j2\pi \frac{lm}{N}} Ev(n)v(l) \tag{5.69}$$

$$= N^{-1} \sum_{n=0}^{N-1} \sum_{l=0}^{N-1} e^{-j2\pi \frac{(nk-lm)}{N}} \sigma_v^2 \delta_{n,l}$$

$$= \frac{\sigma_v^2}{N} \sum_{n=0}^{N-1} e^{-j2\pi n \frac{(k-m)}{N}}$$

$$= \frac{\sigma_v^2}{N} (N \delta_{k,m})$$

$$= \sigma_v^2 \delta_{k,m} \tag{5.70}$$

and so the DFT domain noise is also white, with variance σ_v^2, and being a linear combination of gaussian quantities it is also gaussian.

The analysis of the effect of this noise on the vector output (5.63) of the quadrature filter system is complicated, as noted in [78], by the nonlinearities in (5.61) and (5.62). However, the authors of [78] justify a gaussian approximation by virtue of the Central Limit

Theorem, and observe that this approximation is adequate in practice.

Accordingly, each component of the vector $\mathbf{g}(x,y)$ will be assumed to be subject to additive white gaussian noise. Normalisation of the filter energies further admits the assumption that the noise has variance σ_v^2, equal to the variance of the corrupting noise in the image.

A measure of the error in the orientation estimate, introduced by Willsky[144] and used in [78], is given as a function of the angular error,

$$J(\phi_\varepsilon^{x,y}) = 1 - \cos\phi_\varepsilon^{x,y} \tag{5.71}$$

where

$$\phi_\varepsilon^{x,y} = \phi^{x,y} - \frac{1}{2}\arg\left[\mathbf{g}(x,y)\right] \tag{5.72}$$

is the angular error and $\phi^{x,y}$ is the true orientation of the image feature. However, in regions of near constant gray level, $\phi_\varepsilon^{x,y}$ will be strongly affected by noise, whereas in fact the length $|\mathbf{g}(x,y)|$ of the orientation vector may be close to zero. In this case the error $\phi_\varepsilon^{x,y}$ in its argument is not as serious as the same angular error in oriented regions with larger $|\mathbf{g}(x,y)|$. Accordingly, the error measures used here are the squared vector error

$$sve(x,y) = |\mathbf{g}(x,y) - \mathbf{k}(x,y)|^2 , \tag{5.73}$$

where $\mathbf{k}(x,y)$ is the estimate derived from the uncorrupted image, and its sample mean

$$P_{sserve} = 4^{-\gamma}\sum_{x=0}^{2^\gamma-1}\sum_{y=0}^{2^\gamma-1} sve(x,y) . \tag{5.74}$$

A logarithmic signal-to-noise ratio may also be defined, as

$$SNR_{\phi} = 10 \log_{10} \left[\frac{\mathrm{var}(\mathbf{k}(x, y))}{P_{\text{noise}}} \right] \text{ dB} \qquad (5.75)$$

where $\mathrm{var}(\mathbf{k}(x, y))$ is the sample variance (sample mean squared length) of the vector $\mathbf{k}(x, y)$.

5.5.1 Examples of Noisy Orientation Estimates

Figures P32 — P37 show the effect of noise on the orientation estimate of (5.63). A key is given in table 5.1. The two components of the orientation vector $\mathbf{g}(x, y)$ (which are denoted as the 'cos' and 'sin' components in the table) are displayed. Also displayed is the modulus of the vector, which indicates the strength of oriented features in the vicinity. The orientation estimator was run on the bottom five levels of the quadtree of the noisy image; only the image plane is shown in the figures.

The test image used is the 'blobs' image, both clean and noisy (0dB) versions of which were used as input to the orientation estimator. The range of each displayed quantity was linearly mapped on to the range (0 — 255) of the framestore. Hence for the modulus component (figures P32 and P35) black represents zero and white represents the maximum value attained. For the cos and sin components (P33, P34, P36, P37) the background gray represents zero with white positive and black negative.

Image	Input Image	Modulus	cos component	sin component
'Blobs', clean	P5	P32	P33	P34
'Blobs', 0dB	P8	P35	P36	P37

Table 5.1

Orientation Estimates

Table 5.2 shows the values of SNR_ϕ of (5.75) for the 'girl' and 'blobs' images at input SNRs of 12dB, 0dB and −12dB. The approximate correspondence between the input and orientation SNRs, given the normalised filter energies, supports the contention that the orientation-domain noise may be approximated as additive white gaussian noise of the same variance σ_ϕ^2 as the original corrupting noise.

Image	'Girl'	'Blobs'
Input SNR	SNR_ϕ	
12dB	9.3dB	8.2dB
0dB	−1.5dB	−2.7dB
−12dB	−12.3dB	−13.9dB

Table 5.2

Effect of Input SNR on Orientation SNR

The effect of the input noise is clearly visible as noise in the orientation domain. This latter noise appears subjectively to be fairly evenly distributed spatially, although verification of this hypothesis is difficult due to the nonlinearities mentioned above. Clearly, however, some benefit might be expected from restoration of the noisy

orientation estimates. This will be considered in Chapter 6.

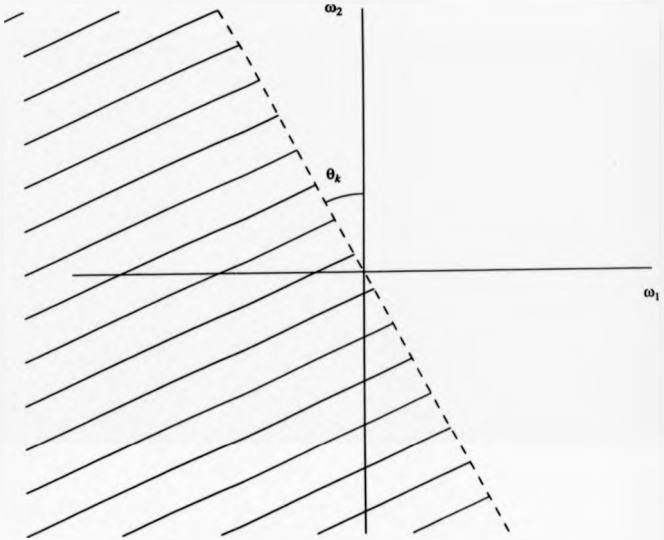


Figure 5.1 — Partition of ω-plane

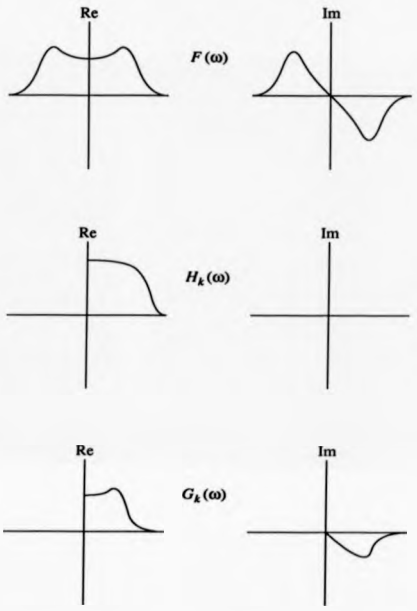
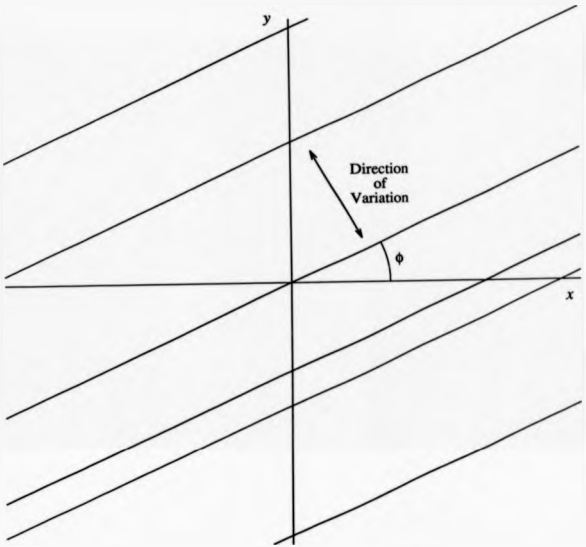


Figure 5.2 — 1-Dimensional Form of Signals in Quadrature Filters



Lines represent contours of equal gray level

Figure 5.3 — Signal of 1-Dimensional Variation

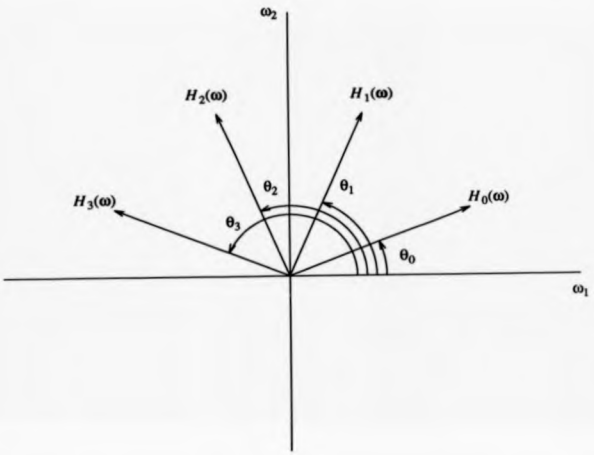


Figure 5.4 — Orientations of Quadrature Filter Pairs in Frequency Domain

CHAPTER 6

QUADTREE ANISOTROPIC ESTIMATION

6.1 Motivation

As noted in section 5.1, visually superior results have been obtained by restoration schemes which utilise anisotropic filtering[75]. The estimated local orientation of image features, such as lines and edges, is used to control the direction in which the anisotropic filter blurs the noisy image.

It is thus possible to blur the noise in a direction parallel to the orientation of the nearby feature, which improves the local signal-to-noise ratio while preserving the definition of the feature.

Such definition is an important component of the subjective assessment of image quality (see section 5.1); an image in which edges are preserved may appear subjectively to be of higher quality than one in which they are blurred, even if the blurred image has a superior signal-to-noise ratio (i.e. lower mean squared error).

This fact constitutes a major disadvantage of conventional Wiener- or Kalman-type image estimators, the lowpass character of which leads to the blurring of features, even when the filter is optimal in terms of the MMSE criterion. The criterion, unfortunately, weights equally errors occurring anywhere in the image regardless of local context, whereas the visual system does not; the latter is far more sensitive to the blurring of edges and rather less sensitive (due to the masking effect — see section 4.4) to noise in the vicinity of such edges.

Quite apart from the difficulty associated with establishing a sufficiently complete description of the fidelity criterion of the visual system is the likely computational complexity or even intractability of the estimator which would be optimal under such a criterion.

Accordingly, those schemes which have attempted to incorporate some degree of adaptation to local image structure have tended to be derived by heuristic arguments and to use more or less *ad hoc* techniques which are justified by their subjective visual results.

This chapter describes the extension of the estimator of Chapters 3 and 4 to include anisotropic processing which is controlled by the orientation detector of Chapter 5. The quadtree vector estimator of Chapter 5 is first used to restore the noisy orientation information.

Since the MMSE criterion is not anisotropic, the anisotropic estimation scheme again departs from the basis of MMSE optimality as did the modifications introduced in Chapter 4. As will be shown, however, the extension to anisotropic processing gives improved visual results. Just as the spatially-variant form of the estimator may, neglecting the addition of the interstitial nodes, be expressed as optimal in terms of the non-contextual and hence spatially-invariant MMSE criterion by the deployment of the signal-equivalent model formulation, a more elaborate and anisotropic signal-equivalent model could be developed, under which the anisotropic estimator would be MMSE optimal.

6.2 Quadtree Restoration of Noisy Orientation Estimates

As noted in section 5.5, each component of the orientation vector $\mathbf{g}(x, y)$ may be considered to be corrupted by additive white gaussian noise of variance σ_v^2. If the quadtree-based model of the form of (2.4) or equivalently (2.6) were found realistically to describe the behaviour of these vector components, then the quadtree estimator of Chapter 3, or its vector form of Chapter 5, might be used to restore the noisy orientation data.

The effectiveness of the model in representing the orientation data is unfortunately not clear. However, it is reasonable to suppose that the orientation vector calculated from an image block of $(2N \times 2N)$ pixels would be close to the vector average of those derived from its four constituent $(N \times N)$ blocks, and in this sense the quadtree model may be appropriate.

The orientation vector field was smoothed using the estimator coefficient values (the e_l of (3.79)) derived from the corrupted image for the case of the spatially-invariant model. This may appear somewhat arbitrary, but it seems appropriate to apply the same amount of smoothing to the orientation data as to the image itself, given that both are subject to the same corrupting noise. This argument applies on any level of the quadtree, and so justifies the use on a given level of the same estimator coefficient for the orientation data as used for the gray-level image data. There is the additional benefit that the coefficients do not then have to be recalculated from the orientation data. A further advantage is that the degree of smoothing applied, and hence spatial spreading induced, is similar to that effected upon the gray-level image data by the estimator proper. Hence the anisotropically filtered bands will be of the same order of width as the residual noise bands which remain in the spatially-variant estimates of section 4.10, and it is precisely those bands which it is desired be smoothed anisotropically.

6.2.1 Examples of Restored Orientation Estimates

Some examples of the restorations of the orientation vector field are shown in figures P38 — P40, as indicated in table 6.1. The input image is the 'blobs' image for input signal-to-noise ratio of 0dB. The restorations may be compared with the noisy estimates of section 5.5.1, indicated in parentheses in table 6.1. Again only the image plane is shown, although the smoothing is performed over the bottom five levels of the quadtree.

Image	Modulus	cos component	sin component
'Blobs', 0dB	P38 (P35)	P39 (P36)	P40 (P37)

Table 6.1

Restored Orientation Estimates

(Unrestored in parentheses)

6.2.2 Mean Squared Vector Error

Table 6.2 shows the effect of the restoration on the mean squared vector error in the orientation estimate. The error is again expressed as the signal-to-noise ratio SNR_{ϕ}.

At the lower input SNRs, the scheme achieves a marked improvement both subjectively and numerically. At the relatively high input SNR of 12dB, the restorations achieve no numerical improvement — this is due to the spreading of the 'edge bands' as a result of the applied smoothing, which introduces an error of magnitude comparable to that due to the noise which is smoothed out in the restoration. This is not a problem, however, since it implies that in the vicinity of an oriented feature, the data will take on the orientation

of that feature rather than its previous noise value.

Image	'Girl'	'Blobs'
Input SNR	SNR_ϕ	
12dB	7.2dB (9.3dB)	8.4dB (8.2dB)
0dB	2.5dB (−1.5dB)	2.2dB (−2.7dB)
−12dB	0.5dB (−12.3dB)	−0.3dB (−13.9dB)

Table 6.2

SNR_ϕ for Restored Orientation Estimates

(Unrestored in parentheses)

6.3 Anisotropic Filtering

In order to incorporate anisotropy into the estimation strategy, a directionally filtered version of the noisy data is computed and made available to the estimator. The form of the combination of this data will be addressed in section 6.4.

Some heuristic design specifications for the anisotropic filtering may be stated as follows:

(i) The anisotropic filter should be of narrow angular bandwidth in order to minimise undesired smoothing in the direction perpendicular to the local orientation estimate.

(ii) It is preferable to use spatially-invariant filters and to interpolate over their outputs (according to the local orientation estimate) in a spatially-variant manner, rather than to recompute the anisotropic filter at each pixel.

(iii) Such interpolation should be unbiased in the sense that a given value of the local orientation should give rise to an anisotropic filter which is oriented in exactly the same direction.

The specifications (i) of narrow angular bandwidth and (ii) interpolation over fixed filters are mutually incompatible. They may, however, be reconciled[75] by iteration of the interpolated filter, provided that the filter is appropriately designed.

6.3.1 Design of the Filter

The filter design is performed in the frequency domain, since the effects of iteration are more easily evaluated than in the spatial domain.

The radial frequency response of the fixed anisotropic filters was chosen as

$$H_r(\rho) = \begin{cases} \frac{1}{2}\sin^2 2\pi\rho \ , & 0 < \rho < 0.25 \ , \ 0.75 < \rho < 1 \\ \frac{1}{2} \ , & 0.25 < \rho < 0.75 \ , \end{cases} \qquad (6.1)$$

iteration of which will tend to sharpen the roll-off at the upper and lower frequencies. $\rho = 1$ is defined here as the maximum (Nyquist) frequency.

Two fixed anisotropic filters are then given by

$$\begin{aligned} H_c(\rho,\theta) &= H_r(\rho)\cos 2(\theta-\theta_0-\tfrac{\pi}{2}) \\ &= -H_r(\rho)\cos 2(\theta-\theta_0) \end{aligned} \qquad (6.2)$$

and

$$H_s(\rho,\theta) = H_r(\rho)\sin 2(\theta-\theta_0-\frac{\pi}{2})$$
$$= -H_r(\rho)\sin 2(\theta-\theta_0) \qquad (6.3)$$

where θ_0 is the offset angle for the orientation vector as given by (5.64). The extra phase term of $\pi/2$ in these equations is included because transformation to the spatial domain rotates the oriented energy of the filters through an angle of $\pi/2$ radians (see the lemma in section 5.4.3, proved as (5.47)).

The interpolation over these two fixed filters is controlled by the local orientation estimate,

$$H_a^{x,y}(\rho,\theta) = \frac{g_c(x,y)}{|\mathbf{g}(x,y)|}H_c(\rho,\theta) + \frac{g_s(x,y)}{|\mathbf{g}(x,y)|}H_s(\rho,\theta) \qquad (6.4)$$

where $g_c(x,y)$ and $g_s(x,y)$ are the components of the orientation vector $\mathbf{g}(x,y)$ (see (5.63)).

The filter is given finally by the addition of two isotropic terms. The first is a term $H_r(\rho)$ as given by (6.1), which converts the cosinusoidal angular response of $H_a^{x,y}(\rho,\theta)$ into a raised cosine. The second is given by

$$H_i(\rho) = 1 - 2H_r(\rho) \ ,$$
$$\rho < 0.25 \ , \qquad (6.5)$$

which gives the filter and its iterations a gain of unity at d.c. and a flat unity-gain passband up to $\rho = 0.75$ along the oriented axis.

The composite filter is then given as

$$H^{x+y}(\rho,\theta) = H_i(\rho) + H_r(\rho) + \frac{g_c(x,y)}{|g(x,y)|}H_c(\rho,\theta) + \frac{g_s(x,y)}{|g(x,y)|}H_s(\rho,\theta) \qquad (6.6)$$

or

$$H^{x+y}(\rho,\theta) = H_p(\rho) + \frac{g_c(x,y)}{|g(x,y)|}H_c(\rho,\theta) + \frac{g_s(x,y)}{|g(x,y)|}H_s(\rho,\theta) \quad , \qquad (6.7)$$

where $H_p(\rho) = H_i(\rho) + H_r(\rho)$. Hence there are in fact three fixed filters which must be implemented.

The filters described above and the interpolation between them may be shown to satisfy specification (iii) above. Assuming the orientation vector at (x,y) to be, from (5.63),

$$g(x,y) = |g(x,y)|\begin{bmatrix}\cos 2(\phi-\theta_0)\\ \sin 2(\phi-\theta_0)\end{bmatrix} \qquad (6.8)$$

where ϕ is the orientation of the local image feature at (x,y), then from (6.4),

$$H_a^{x+y}(\rho,\theta) = -H_r(\rho)\cos 2(\phi-\theta_0)\cos 2(\theta-\theta_0) - H_r(\rho)\sin 2(\phi-\theta_0)\sin 2(\theta-\theta_0)$$

$$= -H_r(\rho)\cos 2(\theta-\phi) \quad . \qquad (6.9)$$

Then for the composite filter, (6.6) yields

$$H^{x+y}(\rho,\theta) = H_i(\rho) + H_r(\rho)\Big[1 - \cos 2(\theta-\phi)\Big] \quad . \qquad (6.10)$$

The effect of the combination is that in the oriented direction, the filter is all-pass (unity gain, zero phase) up to $\rho = 0.75$ after which point a raised-cosine roll-off obtains, whilst in the perpendicular direction a raised-cosine roll-off maintains unity gain at $\rho = 0$ dropping to zero gain at $\rho = 0.25$. These two cases are illustrated in figure 6.1.

6.3.2 Examples of Anisotropically Filtered Images

Noisy images which were filtered using the system developed above are shown in figures P41 to P44. The orientation vector field was estimated from the noisy originals and restored as described in section 6.2. The anisotropic filtering was controlled by the restored orientation estimate as described above. Table 6.3 provides a key to the figures. Only the image plane is displayed; the anisotropic filter was run on the bottom four levels of the quadtree, for which restored orientation estimates are available. Four iterations of the anisotropic filter were performed in order to reduce its angular bandwidth. The spatial filters were confined to masks of (7 x 7) pixels.

Image	Noisy Input	Filtered
'Girl', 12dB	P16	P41
'Girl', 0dB	P7	P42
'Blobs', 12dB	P21	P43
'Blobs', 0dB	P8	P44

Table 6.3

Directionally-Filtered Examples

The filtering clearly achieves the desired effect of smoothing parallel to proximal image features and thus preserving edges. In fact, the edges are even somewhat enhanced by the sidelobes of the filter. This is apparent particularly in the example of figure P43 ('blobs', 12dB) by inspection of the smallest blobs (compare with figure P21).

As the input SNR is reduced, the increased smoothing which is applied to the orientation

estimate causes merging of adjacent edges. This effect gives rise to the fusion of neighbouring blobs which is evident in figure P44.

In the examples of figures P41 and P42, the fine detail of the eyes is smoothed undesirably by the filtering. This results because at the scales of the orientation detection and restoration and the anisotropic filtering, such regions are essentially isotropic and of high bandwidth. The same problem was encountered by Knutsson, Wilson and Granlund[75].

6.4 Anisotropic Extension of the Quadtree Estimator

6.4.1 Form of the Modified Estimator

The estimator of (3.77) may be modified to include an anisotropic component as follows:

$$\hat{s}_{l+1,i',f'} = e_{l+1}\hat{s}_{l,i,j} + (1-e_{l+1})y_{l+1,i',f'} \quad , \quad (l+1,i',f') \in D_{l,i,j} \quad , \tag{6.11}$$

where

$$y_{l+1,i',f'} = ad_{l+1,i',f'} + (1-a)x_{l+1,i',f'} \tag{6.12}$$

and

$$d_{l+1,i',f'} = x_{l+1,i',f'} \otimes h^{i',f'}(\rho,\theta) \tag{6.13}$$

is the directionally-filtered data, where $\otimes$ denotes convolution.

6.4.2 Directed Energy as a Control Parameter

The parameter a of (6.12) controls the combination of the directionally-filtered and 'raw' data prior to the estimation of (6.11). In order that oriented features be preserved, a should be an increasing function of $|g(x,y)|^2$, the oriented energy. Thus in areas which are of high bandwidth but should be considered isotropic (such as the eyes in the 'girl' image), a will be small and little of the directionally-filtered information will be included. (In such areas, directional filtering corresponds to undesirable blurring.)

Thus the parameter a is chosen as

$$a = a^{x,y} = \frac{|g(x,y)|}{\max_{x,y} |g(x,y)|} \quad . \tag{6.14}$$

A more general but computationally expensive scheme might recompute the filter at each point as a function of orientation direction, signal-to-noise ratio for the given quadtree level and the degree of local anisotropy; such a scheme might give improved results, but the computational benefits of the quadtree estimator and edge detector of Chapters 3 and 4 start to be distinctly undermined by the computational burden of the orientation estimation and the directional filtering.

6.5 Results

The modified estimator of (6.11), combined with the quadtree-based edge detector to give the spatially variant form of the estimator of section 4.6, was run on the same noisy images as were used in the production of the 'isotropically' estimated results given in section 4.10.

The scheme now comprises the original optimal estimator of Chapter 3 with the insertion of interstitial nodes (Chapter 4), nonstationarity controlled by the activity detector (Chapter 4), estimation (Chapter 5) and restoration (Chapter 6) of orientation, and the spatially-variant introduction of anisotropically-filtered data into the image estimation (Chapter 6).

6.5.1 Examples of Estimated Images

Examples of images estimated with the anisotropic extension of the scheme are shown in figures P45 — P53. Table 6.4 provides a key to the figures.

Image	Noisy Input	'Isotropic' Estimate	Anisotropic Estimate
'Girl', 12dB	P16	P17	P45
'Girl', 0dB	P7	P18	P46
'Girl', −12dB	P19	P20	P47
'Blobs', 12dB	P21	P22	P48
'Blobs', 0dB	P8	P23	P49
'Blobs', −12dB	P24	P25	P50
'Lake', 12dB	P26	P27	P51
'Lake', 0dB	P28	P29	P52
'Lake', −12dB	P30	P31	P53

Table 6.4

Anisotropically Estimated Examples

Note that differences in film processing seem to have imparted a slightly darker and

lower-contrast appearance to the anisotropically estimated examples than their 'isotropic' counterparts. This is not the case when these images are viewed on a CRT monitor.

6.5.2 Improvement of Signal-to-Noise Ratio

Table 6.5 shows the signal-to-noise ratios of the anisotropically estimated examples, with the corresponding 'isotropic' values in parentheses.

Image	'Girl'	'Blobs'	'Lake'
Input SNR		Output SNR	
12dB	18.8dB (18.8dB)	19.4dB (22.1dB)	17.0dB (17.2dB)
0dB	13.8dB (13.6dB)	14.6dB (15.0dB)	11.8dB (12.0dB)
−12dB	8.8dB (8.5dB)	9.6dB (9.5dB)	7.1dB (7.2dB)

Table 6.5

SNR of Anisotropic Estimates

(Isotropic in parentheses)

Numerically, the anisotropic estimator performs better than the 'isotropic' scheme for certain of the test images and worse for others. The reason for the poorer performance on the 'blobs' image at an input SNR of 12dB is visible in figure P48 as a dark band around each of the objects. A similar light band within each object exists but is rather less visible. This phenomenon is caused by the sidelobes of the anisotropic filter, which produce a dark band (see section 6.3.2, figure P43) around each object on the bottom four levels of the quadtree. When this information is combined as per (6.11) the bands from each of these four scales appear concentrically at the image plane and together form the dark

band observed in figure P48.

6.5.3 Discussion

The anisotropic estimator provides a distinct subjective improvement over the 'isotropic' estimator of Chapter 4, particularly in the case of the natural ('girl' and 'lake') images. Much of the noise which breaks through around edges in the isotropic examples is removed by the anisotropic method, as is clear by inspection of the 0dB examples (compare figures P18 with P46 and P29 with P52). The control exerted by the oriented energy over the incorporation of the anisotropically-filtered data into the information available to the estimator is demonstrated by the retention of such features as the eyes in the 'girl' image and the leaves of the trees in the 'lake' image, despite the considerable degree of anisotropic smoothing applied in oriented regions such as the tree trunks in the 'lake' image.

The restorations at the very low input SNR of −12dB are distinctly acceptable; little or nothing which is visible to the viewer in the noisy originals fails to appear in the restorations.

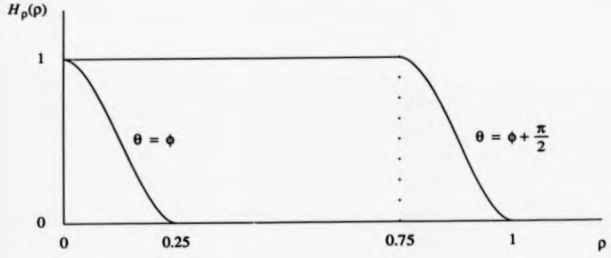


Figure 6.1 — Extremal Radial Frequency Responses of Anisotropic Filter

CHAPTER 7

CONCLUSIONS AND FURTHER WORK

This work concerns the restoration of digital images from corrupted data, using a novel linear estimation strategy based upon a new model for images.

The model was developed in response to the need for more realistic statistical image descriptions, since existing models tend to be deficient in a number of important respects.

Chapter 1 provides an introductory discussion of the salient properties of the principal classes of image model which have been employed to date in various image processing contexts, and contrasts these with some of the characteristics of natural images which are deduced from the structure of the physical world. The application of the models to optimal estimation is examined in relation to the various different criteria of optimality (the 'fidelity' criteria) on the basis of which the estimation strategy may be designed.

A dilemma in optimal image estimation is explored; the simple, highly tractable and commonly used minimum mean squared error (MMSE) criterion is observed to give poor visual results because the fidelity criterion of the human visual system is profoundly different from MMSE. Indeed, the visual system is known to exhibit strong contextual effects in terms for instance of its sensitivity to noise in different image regions and to the blurring of edges, which is a consequence of estimation under shift-invariant optimality criteria such as MMSE. However, the tractability of the MMSE criterion motivates its retention as the basis of optimality, shifting the emphasis in the quest for visually more satisfactory estimation strategies on to the image model. Accordingly, if an image model could be developed which:

(i) is realistic in terms of its representation of the image;

(ii) incorporates desirable features of the more complicated visual optimality criterion, and

(iii) admits of a simple MMSE estimator,

then this model should provide the basis for a much more visually satisfactory estimation scheme than is possible with existing models and simple MMSE estimation. If in addition the model structure were relatively uncomplicated, these advantages might be realised without undue computational burden. This is the key motivation for the present work.

Among the common attributes of image models are causality and stationarity; these paradigms are examined in Chapter 2.

Causality describes the presence of a governing 'time' direction in the data, and implies that at any given instant some of the signal data (that segment which is supported by the independent variable at indices greater than the present index) is unavailable to the system. Causal models have been used to represent image data particularly when that data is presented to the processing system in raster-scanned format, due to both their effectiveness in one-dimensional signal processing where the independent variable is 'real' time, and the unrealisability of noncausal ('anticipative') physical systems.

This work argues, however, that causality is an inappropriate attribute for image models and represents at best merely a model of one particular image acquisition method. The notion of causality is not easily extended to the case where the data is a function of two or more independent variable indices, and is particularly incongruous when those indices have positional or spatial rather than temporal significance, since in that case there exists

in the data no *a priori* dominant ordering.

However, many causal models do admit of recursive expression, and associated linear filtering operators may often be coerced into the same form with a substantial saving in computational effort by comparison with noncausal methods. This feature of causal schemes is highly attractive and represents a major advantage over their noncausal counterparts which it is desired be retained in the present work despite the rejection of causality as a signal attribute.

Another common feature of image models has been stationarity. This implies that the model parameters are shift-invariant, and hence that the model generates homogeneous images. It is almost trivially obvious by inspection of a few examples, however, that natural images are far from homogeneous and tend to contain distinct objects with more or less uniform interiors and with well-defined and locally specifically-oriented boundaries. This work argues therefore that any attempt to devise a more realistic image model must include the capacity for spatial variation (nonstationarity) of the model parameters. The potential attendant disadvantage is the increased complexity and computational burden usually associated with nonstationary processing.

Chapter 2 considers the various approaches to the problem which have been adopted; these include the so-called 'multiple-model' methods. The work argues that these methods, in which a number of stationary models are deployed to represent different types of image region and are combined in a spatially-variant manner to yield the composite image model, are deficient by virtue both of their assumption of local homogeneity and their inability accurately to represent image regions of arbitrary structure if that structure does not conform to any of the limited number of available models. An argument is made for 'full' nonstationarity in the sense of a single general model which is

capable of representing any image structure with appropriate variation of its parameters.

Another property of natural images which it is desired be incorporated into the new model is scale invariance. This notion as it applies to natural images derives from the same property exhibited by the physical world and is a consequence of the particulate and cohesive nature of matter. Natural images demonstrate similar structure and organisation over a range of scales, whereas any conventional image model of given order must necessarily be best matched to a particular organisational scale related to that order.

The new image model is presented in Chapter 2. It is defined on the so-called 'quadtree', which is a multiresolution data structure for image representation. The model characterises an image as a sample of a random process which is generated by a number of component gaussian processes operating at different levels of the quadtree structure. The multiresolution facility makes explicit the behaviour of the model at each of a range of scales in the generated image; the model behaves in an identical but entirely independent fashion at each scale as a result of its definition on the quadtree.

The model generates an image which is noncausal, but the model itself is causal in the context of the level index of the quadtree, which may be regarded as a third index of higher order than the positional indices on each level and may be considered to represent a third dimension orthogonal to the spatial plane. Thus a noncausal two-dimensional signal is generated by a one-dimensional model causal in this third dimension; the system generalises to any dimension N and generates a noncausal signal by confining the model causality to an extra dimension N+1. The computational advantages of the one-dimensional causal paradigm are retained but without the restriction to causality, or to any given dimension, of the generated signal.

The model is 'fully' nonstationary in the sense that its parameters are able to vary arbitrarily over their unbounded range with spatial position on a given quadtree level. Furthermore, the parameters are able to vary over the same range between levels for a given image position. This corresponds to the scale invariance and scale independence attributes described above; these two types of nonstationarity allow complete variation of the model activity within a conceptual three-dimensional volume occupied by the quadtree, in which two dimensions represent spatial position and the third (the 'vertical') represents scale. As noted above, the model is causal in the 'scale' index.

Chapter 2 presents an analysis of the statistical properties of images generated by the model for the spatially-invariant case. The correlation properties of the signal are derived, and the system is shown to be diagonalised or decorrelated by the Hadamard and Haar discrete transforms.

Chapter 3 deals with the application of optimal estimation to the data structure in the context of the restoration of images corrupted by additive white gaussian noise, which is first justified as a suitable model for noise degradation. The model for the noisy image is established, and the formation of a quadtree of the noisy data is described and its correlation properties are derived.

Optimal (MMSE) estimates of the signal value at any quadtree node are derived both for the constrained case in which the available data lies only in the projection of the node on to the image, and in the general case where the data set consists of the entire quadtree. The general MMSE estimator is shown to reduce to a recursive form which constitutes a one-dimensional scalar Kalman filter operating on the quadtree and causal in the level index like the image model. The constrained estimate at the uppermost ('root') node is shown to constitute the MMSE initial condition for the recursion, and the optimal estima-

tor of the model innovations process is combined with the optimal node-value estimator to yield a realisation of the scheme which involves only 5/4 multiplications per node on average. This corresponds to 5/3 multiplications per image pixel, a figure far lower than that achieved by any other noncausal scheme and considerably lower even than the burden associated with the simplest of causal methods.

An alternative derivation of the estimator is presented in Appendix 1 in terms of the classical Yule-Walker or 'normal' equations for multivariate MMSE estimation, the results then being coerced into the recursive form by a method analogous to the recursion of Levinson[84][110]. A form of correlation matrix for the noisy data is presented and is shown to be invertible by a recursion to yield a tridiagonal inverse.

Appendix 2 carries forward the Kalman filtering development of section 3.3, and derives the Kalman whitening filter and the Kalman innovations process for the noisy data. The innovations are shown to be proportional simply to the differences between the averages on successive levels of the data quadtree.

The MMSE estimator is derived in Appendix 2 in a manner similar to that of section 3.3.2, noting that there are between the two cases a number of significant differences in the signal and 'noise' statistics. The 'noise' for the purposes of the data quadtree is non-white and is correlated with the signal. However, it is shown that the differences between the 'next' datum and the 'present' and 'next' estimates respectively may each be expressed as a non-white signal which is a linear combination of the Kalman innovations of the data, with coefficient values which are such that a recursive expression of the estimator is forthcoming.

The final part of Appendix 2 deals with the Cholesky factorisation[110] of the data corre-

lation matrix and its inverse. The inverse of this matrix is shown to possess a causal/anticausal factorisation with the causal factor given by the Kalman whitening filter and the anticausal by its transpose (see [110]). The Kalman innovations filter of the data is obtained by recursive inversion of the whitening filter and is shown to constitute the causal factor in a similar Cholesky decomposition of the correlation matrix.

The estimator is shown in Chapter 3 to be highly amenable to parallel computation by an array (ideally a quadtree) of processing elements.

Chapter 3 considers also the estimation of the optimal estimator coefficients from the noisy image data, and shows that the coefficients thus estimated obey the so-called F probability density function[8]. Appendix 3 provides a full derivation of the result.

Initial estimation results are presented for the case of spatially-invariant processing, and two major problems are observed in the results. The first is the pronounced block structure of the estimated images, which is a consequence of the spatial isolation between nodes in the quadtree. The second is the blurring of edges which follows from the shift invariance of the estimation.

Chapter 4 addresses both of these problems.

Firstly, the problem of blocking in the estimate is considered. In order to reduce this effect, some kind of spatial filtering is required. However, it is desired that the quadtree structure not be compromised to any great extent, with spatial operations having been avoided thus far. Indeed, this is precisely the reason for the very low computational burden associated with the estimator, and so spatial filtering is somewhat undesirable. The solution, presented in Chapter 4, is the introduction of additional ('interstitial') nodes into

the quadtree structure. These nodes form, on each level, a lattice which is displaced from the existing lattice by half of the spatial sampling interval in each positional coordinate. The interstitial nodes are shown collectively to straddle all of the block boundaries in the quadtree. The computation is minimally modified: in the formation of the quadtree of noisy data, the interstitial nodes are assigned the average value of their four 'child' nodes, while in the estimation phase the 'father' of a given node is taken as the average of the existing father and an interstitial node. No multiplications are demanded by the scheme, as would be the case for conventional spatial filtering, but the block effects are entirely eliminated. This approach is therefore adopted whenever data are passed up or down the quadtree.

The second problem with the estimation of Chapter 3 is its spatial invariance. Nonstationary operation is vital if edges are to be preserved and not blurred, and with the removal of the blocking effect it is apparent that the spatially-invariant estimates are indeed blurred. Chapter 4 introduces the spatially-variant or nonstationary form of the estimator. This involves the modification of the stationary method on the basis of contextual information extracted from the image by an 'activity' detection scheme defined on the quadtree.

The design is motivated by a number of factors. These include

(i) the scale invariance of the image model and the facility for fully nonstationary processing at all quadtree levels;

(ii) the notion that edges in images are represented at a number of contiguous scales of resolution;

(iii) the possibility of effecting some degree of noise rejection by combining information from different scales, since noise at the different scales may be expected to be

largely uncorrelated, and

(iv) the observation that the spatially-invariant estimator requires a minimal operation count by virtue of its definition on the quadtree and its avoidance of spatial operators.

Two indices of local image activity are presented. The first, 'simple', index is just the normalised energy of the link between a node and its 'father'; the second is a propagated activity index, defined recursively as a weighted geometric mean (WGM) of the same quantity at the father and the 'simple' index at the current node. The weighting is a function of the input signal-to-noise ratio, and the WGM achieves noise rejection because of its weighted-'AND' character. Activity represented in the 'simple' index which is due to noise tends to be reduced in the propagated index because there is little positional correlation between the energy on each quadtree level. In the case of a genuine edge, however, the energy is to be found at the same spatial position over a number of scales, and the 'AND' inherent in the WGM reinforces the energy at the appropriate location in the propagated index.

The activity detector inherits many of the advantageous characteristics of the estimator itself. These include scale invariance and light computational burden resulting from the lack of a spatial operator as used by many existing methods. The notion of activity being represented at multiple scales fits easily into the design, whereas operator-based edge detectors which utilise multiple-scale information are compelled to run a number of different spatial filters in order to create the 'scale space'. The present scheme, like the estimator, is amenable to parallel computation. The ability of the scheme to localise activity such as edges in the presence of noise is observed to be limited by an uncertainty.

The propagated activity index is used to modify the estimation in such a way that active

regions such as edges in the image are subject to reduced smoothing by comparison with the spatially-invariant case, whereas regions of relatively constant luminance are smoothed to a degree which is greater than in that case. The form of the modification is such that quadtree nodes which exhibit activity typical of their level retain the original stationary estimator coefficient. This is appropriate because that coefficient results in the MMSE estimate based on the global statistical properties (i.e. the activity) of the image.

Chapter 4 presents the 'signal-equivalent' spatially-variant form of the image model, for which the modified nonstationary estimator represents the MMSE solution. The parameters of that model are derived in 'backward' fashion from the coefficients of the modified estimator.

Results are presented which demonstrate the operation of the activity detector and the effect of its output (the propagated index) on the estimator coefficient. The degree of noise rejection achieved is apparent.

Chapter 4 also presents estimation results for the full implementation of the nonstationary estimator. The system achieves results which are subjectively markedly improved by comparison with the stationary scheme. The numerical improvement, in terms of signal-to-noise ratio (SNR), is very respectable, with the 'girl' image improved by 13.6dB and the 'blobs' image by 15.0dB for input SNRs of 0dB. This figure represents a performance superior to that of other systems and considerably in excess of the 8.3dB and 9.3dB achieved by the schemes of [156] and [61] respectively for a 0dB input.

As a result of the nonstationary processing designed to preserve edges, the system allows a 'band' of noise to remain in edge regions. It is observed that the width of this band and the characteristics of the noise therein depend on the input SNR and that these quantities

are governed by the uncertainty which affects the activity detector as noted above.

Chapter 5 deals with the extension of the nonstationary estimator to the case where the image data is vector-valued, as for example in restorations of colour or multispectral images, or of texture descriptor fields. Another application area is the restoration of orientation data computed from noisy images; anisotropic filtering has been used with some success in image restoration by Knutsson, Wilson and Granlund[75] and the method requires data relating to local image orientation as a control input.

The extension to vector data is relatively straightforward. Chapter 5 introduces vector versions of the model, activity detector and estimator. Each component of the vector image may be estimated using a different set of coefficients, in which case the estimator is optimal for each component, or if it is desired that the balance between components be preserved, as would probably be the case in colour or multispectral applications, a single solution may be applied for all components, implying a scheme which is optimal given the constraint.

Chapter 5 emphasises the importance in vision of oriented features, which motivates the development of anisotropic filtering and thus the restoration of orientation data from noisy observations. The initial extraction of orientation information is considered, and the design (following the work of Hans Knutsson[79][80]) of a quadrature filter set for this purpose is presented. The hypothesis that similar mechanisms are utilised in vision is supported by evidence from physiological experiments, chiefly the work of Pollen and Ronner[115][116]. The output of the orientation detectors is in the format of the 'double-angle' vector representation used by Knutsson, Wilson and Granlund[75][78]. The analysis of its susceptibility to noise is noted to be a difficult analytical procedure due to the presence of nonlinearities in the estimation, and simplifying assumptions are used in

order that the orientation-domain noise may be modelled as additive white gaussian. Results are presented which show the effect of noise on the orientation estimates. The noise in the estimate is quantified in terms of a mean squared vector error, which may be expressed in terms of the familiar signal-to-noise ratio by defining the 'signal' to be the orientation field as estimated from noiseless images.

Chapter 6 describes the subsequent utilisation of the orientation vector field. Firstly, the noisy orientation information is restored and results are presented illustrating the restorations. The system uses the spatially-invariant form of the quadtree vector estimator. The numerical improvement is again defined in terms of the mean squared vector error and its associated signal-to-noise ratio.

The design of the anisotropic filters which are to be controlled by the orientation data is presented. Three fixed filters are used, their outputs being interpolated according to the restored local orientation data. The anisotropic filtering is iterated in order to reduce the angular bandwidth and thus to prevent unwanted smoothing in a direction perpendicular to the local feature. Examples of the output of the anisotropic filtering are presented and discussed.

Chapter 6 then treats the extension of the nonstationary estimator to include anisotropic processing. The estimator structure and its definition on the quadtree remain unchanged, but now the data in the quadtree (previously the averages generated from the noisy image) is mixed in regions of high anisotropic energy such as edges with the anisotropically-filtered data. The mixture is controlled by the local modulus of the orientation vector so that in strongly anisotropic regions, the data is almost exclusively the output of the anisotropic filter, while in smooth or isotropic regions the data is little altered. This means that in image localities which are essentially isotropic at the scale

concerned but possibly contain important detail (the eyes in the 'girl' image being a good example) the noisy data is not mixed with the filtered (which in these regions corresponds merely to an undesirable degree of blurring).

Results for the anisotropic extension of the estimator are presented and compared to the results obtained in Chapter 4. While there is generally no numerical improvement over the latter results, the subjective visual improvement is considerably more marked. The 'noise bands' which remained in the vicinity of edges are now much less prominent, while the edges themselves are preserved (and even possibly somewhat enhanced by the sidelobes of the anisotropic filter).

Whilst the extension of the scheme to include anisotropic processing provides a distinct improvement in the subjective quality of the restorations particularly in the vicinity of edges, such processing is not as well matched to the structure of the quadtree as are the other components of the method, and the computational effort is greatly in excess of that required by the unextended scheme. This is a consequence of the need for extraction and restoration of the orientation data and the anisotropic filtering itself.

In the absence of special-purpose hardware for these tasks, the question must arise of whether the improvement justifies the additional computation. Whilst the unextended method might certainly be implemented at real-time data rates on a quadtree of fairly simple processors, the considerable additional burden associated with the anisotropic extension makes real-time implementation of the extended method appear rather less feasible. Further work might profitably examine the possibility of improving the method by some less complicated extension, although it is difficult to envisage a significantly more computationally efficient system, particularly for the orientation estimation, which would nevertheless meet the requisite criteria. However, Watson[141], Watson and Ahu-

mada[140] and Wilson and Calway[145] have investigated multiresolution (pyramidal) data structures in which orientation as well as scale serves as an index. It may be that an image model and estimator based on such a structure and capturing explicitly the notions of orientation and scale might provide a solution to this problem. It remains to be seen whether a more computationally efficient approach, with subjective performance comparable to that of the extended method used in this work, is possible.

Of considerable utility would be the development of optimal estimation strategies for multiresolution signal models more general than the simple quadtree model presented herein. Unfortunately it is easily shown that for a more general case (even that given by the inclusion of the interstitial nodes in the present scheme) the optimal estimator ceases to possess the efficient recursive structure of the quadtree estimator. Nevertheless, just as for the causal Kalman filtering methods of Woods et al[156][157], it is possible to envisage a 'reduced update' form of the estimator which preserves the recursive structure, but in which the strict rules of ancestry of the quadtree are relaxed to accommodate a father:child relation in the pyramid of M:N instead of the present 1:4. As in the case of the reduced update Kalman filter (RUKF) it is unlikely that the loss of optimality would be accompanied by a significant reduction in the improvement obtained. This is strongly suggested by the results achieved here with the inclusion of the interstitial nodes.

APPENDIX 1

DERIVATION OF THE ESTIMATOR BY A LEVINSON-TYPE RECURSION

Throughout this appendix, the notation $F_k = R_{xx}(k, m, n; k, m, n)$ will be used. This is the (spatially invariant) expected energy $E x^2_{k,m,n}$ of a node $x_{k,m,n}$ on level k of the average-value data quadtree.

The estimator $\hat{s}_{l,i,j}$ may be described in terms of a causal, nonrecursive linear combination of the noisy ancestors $x_{k,p,q}$, $(k, p, q) \in A_{l,i,j}$ of the node (l, i, j),

$$\hat{s}_{l,i,j} = \sum_{\substack{(k,p,q) \\ \in A_{l,i,j}}} \lambda^l_k x_{k,p,q} \tag{A1.1}$$

The orthogonality principle gives the solution vector Λ_l as

$$\Lambda_l = R_l^{-1} r_l \tag{A1.2}$$

where

$$R_l = \begin{bmatrix} F_0 & F_0 & F_0 & \cdots & F_0 & F_0 \\ F_0 & F_1 & F_1 & \cdots & F_1 & F_1 \\ F_0 & F_1 & F_2 & \cdots & F_2 & F_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ F_0 & F_1 & F_2 & \cdots & F_{l-1} & F_{l-1} \\ F_0 & F_1 & F_2 & \cdots & F_{l-1} & F_l \end{bmatrix} \tag{A1.3}$$

is the data correlation matrix,

$$\Lambda_l = \begin{bmatrix} \lambda^l_0 \\ \lambda^l_1 \\ \cdots \\ \lambda^l_{l-1} \\ \lambda^l_l \end{bmatrix} \tag{A1.4}$$

is the coefficient vector, and

$$\mathbf{r}_l = \begin{bmatrix} R_{xx}(0,0,0;l,i,j) \\ \cdots \\ R_{xx}(k,p,q;l,i,j) \\ \cdots \\ R_{xx}(l,i,j;l,i,j) \end{bmatrix} \qquad \text{(A1.5)}$$

where $(k,p,q) \in A_{l,i,j}$. (A1.3) follows from the middle equation of (3.50) because $R_{xx}(l,i,j;k,p,q) = R_{xx}(l,i,j;l,i,j) = F_l$ for $(k,p,q) \in D_{l,i,j}$.

Equation (A1.2) is the matrix form of the *normal* , or *Yule - Walker* , equations[86][110] for this realisation of the estimator. The solution of this system involves the inversion of the data correlation matrix R_l.

Theorem

The structure of R_l is such that its inverse is tridiagonal and may be calculated by recursion on l.

Proof

From (A1.3), the matrix R_{l+1} is given by

$$R_{l+1} = \begin{bmatrix} F_0 & F_0 & \cdots & F_0 & F_0 \\ F_0 & F_1 & \cdots & F_1 & F_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ F_0 & F_1 & \cdots & F_l & F_l \\ F_0 & F_1 & \cdots & F_l & F_{l+1} \end{bmatrix} \qquad \text{(A1.6)}$$

and its determinant Δ_{l+1} is given by

$$\Delta_{l+1} = (F_{l+1} - F_l)\,\Delta_l \quad . \tag{A1.7}$$

The cofactors in the augmenting row and column of (A1.6) are given by

$$c_k = \mathrm{cof}(F_k) = \begin{cases} 0 \ , & k < l \ , \\ -\Delta_l \ , & k = l \ , \\ \Delta_l \ , & k = l+1 \ . \end{cases} \tag{A1.8}$$

Iteration of (A1.7) and (A1.8) gives R_l^{-1} as

$$R_l^{-1} = \begin{vmatrix} \frac{F_1}{F_0(F_1-F_0)} & \frac{-1}{F_1-F_0} & \cdots & 0 & 0 \\ \frac{-1}{F_1-F_0} & \frac{F_2-F_0}{(F_2-F_1)(F_1-F_0)} & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \frac{F_l-F_{l-2}}{(F_l-F_{l-1})(F_{l-1}-F_{l-2})} & \frac{-1}{F_l-F_{l-1}} \\ 0 & 0 & \cdots & \frac{-1}{F_l-F_{l-1}} & \frac{1}{F_l-F_{l-1}} \end{vmatrix} \tag{A1.9}$$

and the proof is complete.

Substituting (3.49) and (3.50) into (A1.9) and the result into (A1.2) yields, after extensive manipulation, the solution for the optimal MMSE estimator coefficients λ_k^l as

$$\lambda_k^l = \frac{4^{\gamma-k}\beta_k^2}{Z_k\,Z_{k+1}}\,Z_{l+1} \quad ,$$

$$0 \le k \le l \quad , \tag{A1.10}$$

where

$$Z_m = 4^{Y-m}\beta_m^2 + 4^{Y-m-1}\beta_{m+1}^2 + \cdots + 4\beta_{Y-1}^2 + \beta_Y^2 + \sigma_v^2 \quad , \qquad \text{(A1.11)}$$

and the solution is complete.

The estimator $\hat{s}_{l,i,j}$ of $s_{l,i,j}$ has been expressed in terms of the noisy average values $x_{k,p,q}$, $(k,p,q) \in A_{l,i,j}$ at the ancestors of the node (l,i,j). The system of (A1.1) represents a time-varying causal filter of increasing order l with coefficients λ_k^l, $0 \le k \le l$.

The algorithm of Levinson[84][110] allows the coefficients of a predictor of order $N+1$ to be expressed in terms of those of the predictor of order N of the same signal. By analogy it is possible in the present case to express the coefficients λ_k^{l+1} of the nonrecursive estimator of order $l+1$ in terms of those λ_k^l of the nonrecursive estimator of order l.

From (A1.10) it is clear that

$$\lambda_k^{l+1} = \left[\frac{Z_{l+2}}{Z_{l+1}}\right] \lambda_k^l \quad ,$$

$$0 \le k \le l \quad , \qquad \text{(A1.12)}$$

and

$$\lambda_{l+1}^{l+1} = \frac{4^{Y-l-1}\beta_{l+1}^2}{Z_{l+1}} \quad . \qquad \text{(A1.13)}$$

In order to obtain a recursive structure for the estimator of the form

$$\hat{s}_{l+1,\,i',\,f'} \;=\; e_i^{l+1}\,\hat{s}_{l,\,i,\,j} \;+\; e_{l+1}^{[e]}\,x_{l+1,\,i',\,f'} \quad ,$$

$$(l+1,\,i',\,f') \in D_{l,\,i,\,j} \quad , \tag{A1.14}$$

it is required that

$$e_{l+1}^{l+1} = \lambda_{l+1}^{l+1} \quad ,$$

$$0 \le l \le Y-1 \quad , \tag{A1.15}$$

and

$$e_i^{l+1} = \frac{\lambda_k^{l+1}}{\lambda_k^l} \quad ,$$

$$0 \le l \le Y-1 \quad ,$$

$$0 \le k \le l \quad . \tag{A1.16}$$

From (A1.12) it is apparent that

$$e_i^{l+1} = \frac{\lambda_k^{l+1}}{\lambda_k^l} = \frac{Z_{l+2}}{Z_{l+1}} \tag{A1.17}$$

independent of k, and so (A1.15) and (A1.16) are satisfied (see (3.78) and (A1.11)) by

$$e_i^{l+1} = \frac{Z_{l+2}}{Z_{l+1}} = e_{l+1} \quad ,$$

$$0 \le l \le Y-1 \quad , \tag{A1.18}$$

and

$$\varepsilon_{l+1}^{[+]} = \lambda_{l+1}^{[+]} = 1 - e_{l+1} \ ,$$

$$0 \le l \le Y-1 \ , \qquad \text{(A1.19)}$$

and so (A1.14) is equivalent to (3.77).

APPENDIX 2

THE KALMAN INNOVATIONS OF THE DATA AND THE OPTIMAL ESTIMATOR

A2.1 The Kalman Whitening Filter

The Kalman innovations signal $i_x(n)$ is derived from the data $x(n)$ by the linear combination of (3.10), and is by definition orthonormal as indicated by (3.9). The system of coefficients γ_k^n of the combination is known as the Kalman whitening filter[110] of $x(n)$ (see section 3.3),

$$i_x(n) = \sum_{k=0}^{n} \gamma_k^n x(k) \quad . \tag{A2.1}$$

The system of equations given by the variation of n in (A2.1) over the range $0 \le n \le l$ may be written in matrix form as

$$\mathbf{i}_x^l = \Gamma_l \mathbf{x}^l \tag{A2.2}$$

where

$$\mathbf{i}_x^l = \begin{bmatrix} i_x(0) \\ i_x(1) \\ \cdots \\ i_x(l-1) \\ i_x(l) \end{bmatrix} \quad , \tag{A2.3}$$

$$\mathbf{x}^l = \begin{bmatrix} x(0) \\ x(1) \\ \cdots \\ x(l-1) \\ x(l) \end{bmatrix} \quad , \tag{A2.4}$$

and

$$\Gamma_l = \begin{pmatrix} \gamma_0^0 & 0 & \cdots & 0 & 0 \\ \gamma_0^1 & \gamma_1^1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_0^{l-1} & \gamma_1^{l-1} & \cdots & \gamma_{l-1}^{l-1} & 0 \\ \gamma_0^l & \gamma_1^l & \cdots & \gamma_{l-1}^l & \gamma_l^l \end{pmatrix} \tag{A2.5}$$

is the Kalman whitening filter, which is lower-triangular.

The orthonormality of the innovations process may be expressed (see (3.9)) as

$$E i_x(m) i_x(n) = \delta_{m,n} \tag{A2.6}$$

or, in matrix form,

$$R_{i_x}^l = E \mathbf{i}_x^l (\mathbf{i}_x^l)^T = \mathbf{I}_l \tag{A2.7}$$

where $\mathbf{I}_l$ is the identity matrix of appropriate order.

Substituting (A2.2) in (A2.7) gives

$$\Gamma_l R_l \Gamma_l^T = \mathbf{I}_l \tag{A2.8}$$

where R_l is the data correlation matrix (see Appendix 1).

Hence for $l = 0$,

$$\gamma_0^0 F_0 \gamma_0^0 = 1 \tag{A2.9}$$

or

$$(\gamma_0^0)^2 = \frac{1}{F_0} \quad . \tag{A2.10}$$

Then for $l = 1$,

$$\begin{bmatrix} \gamma_0^0 & 0 \\ \gamma_0^1 & \gamma_1^1 \end{bmatrix} \begin{bmatrix} F_0 & F_0 \\ F_0 & F_1 \end{bmatrix} \begin{bmatrix} \gamma_0^0 & \gamma_0^1 \\ 0 & \gamma_1^1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{A2.11}$$

which yields, with γ_0^0 determined as above,

$$(\gamma_1^1)^2 = \frac{1}{F_1 - F_0} \tag{A2.12}$$

and

$$\gamma_0^1 = -\gamma_1^1 \quad . \tag{A2.13}$$

Recursion on l gives the coefficients γ_k^l as

$$(\gamma_l^l)^2 = \frac{1}{F_l - F_{l-1}} \quad , \tag{A2.14}$$

$$\gamma_{l-1}^l = -\gamma_l^l \tag{A2.15}$$

and

$$\gamma_k^l = 0 \quad ,$$

$$0 \le k \le l-2 \quad , \tag{A2.16}$$

and so the Kalman whitening filter Γ_l is given by

$$\Gamma_l = \begin{vmatrix} \frac{1}{\sqrt{F_0}} & 0 & \cdots & 0 & 0 \\ \frac{-1}{\sqrt{F_1 - F_0}} & \frac{1}{\sqrt{F_1 - F_0}} & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{F_{l-1} - F_{l-2}}} & 0 \\ 0 & 0 & \cdots & \frac{-1}{\sqrt{F_l - F_{l-1}}} & \frac{1}{\sqrt{F_l - F_{l-1}}} \end{vmatrix} \quad , \text{(A2.17)}$$

and the innovations signal by

$$i_x(m) = \begin{cases} \frac{x(0)}{\sqrt{F_0}} = \gamma_0 x(0) = \gamma_0 (x(0) - x(-1)) \text{ if } x(-1) = 0 \\ \frac{x(m) - x(m-1)}{\sqrt{F_m - F_{m-1}}} = \gamma_m (x(m) - x(m-1)) \quad . \end{cases} \quad \text{(A2.18)}$$

Thus the Kalman innovations of the data are proportional simply to the differences between the averages on successive levels of the quadtree, and the lower-triangular matrix which represents the Kalman whitening filter is sparse, having nonzero elements only on the leading diagonal and the adjacent diagonal.

A2.2 Derivation of the Estimator by Analogy with Section 3.3.2

Note that in this section, the signal innovations $w(.)$ are as intended in section 3.3.2 rather than as in section 2.3.

The orthogonality relation for the estimate $\hat{s}(n)$ of $s(n)$ from the data $x(m)$, $0 \le m \le n$, may be written as

$$E(s(n) - \hat{s}(n))\, x(m) = 0 \quad ,$$

$$0 \le m \le n \quad . \tag{A2.19}$$

Hence

$$E(x(n) - \hat{s}(n))\, x(m) = E(v(n) + s(n) - \hat{s}(n))\, x(m)$$

$$= E\, v(n)\, x(m)$$

$$= E\, v(n)\, v(m) \quad ,$$

$$0 \le m \le n \quad , \tag{A2.20}$$

since $E\, v(n)\, s(m) = 0$ for $0 \le m \le n$.

Changing n to $n+1$ in (A2.20) gives

$$E(x(n+1) - \hat{s}(n+1))\, x(m) = E\, v(n+1)\, v(m) \quad ,$$

$$0 \le m \le n+1 \quad . \tag{A2.21}$$

Now

$$E(x(n+1) - \hat{s}(n))\, x(m) = E(s(n) - \hat{s}(n) + w(n+1) + v(n+1))\, x(m)$$

$$= E(w(n+1) + v(n+1))\, x(m)$$

$$= E(w(n+1) + v(n+1))\, v(m) \quad ,$$

$$0 \le m \le n \quad , \tag{A2.22}$$

since $E\, w(n+1)\, s(m) = E\, v(n+1)\, s(m) = 0$ for $0 \le m \le n$.

In summary,

$$E\,(x\,(n+1)-\hat{s}\,(n+1))\,x\,(m)\;=\;E\,v\,(n+1)\,v\,(m)\quad,$$

$$0\le m\le n+1\quad, \tag{A2.23}$$

and

$$E\,(x\,(n+1)-\hat{s}\,(n))\,x\,(m)\;=\;E\,(v\,(n+1)+w\,(n+1))\,v\,(m)\quad,$$

$$0\le m\le n\quad. \tag{A2.24}$$

Now the innovations signal is given by (A2.18) as

$$i_x(m)\;=\;\gamma_m\,(x\,(m)-x\,(m-1)) \tag{A2.25}$$

and so

$$E\,(x\,(n+1)-\hat{s}\,(n+1))\,i_x(m)\;=\;\gamma_m\,E\,(x\,(n+1)-\hat{s}\,(n+1))\,(x\,(m)-x\,(m-1))$$

$$=\;\gamma_m\,E\,v\,(n+1)(v\,(m)-v\,(m-1))\quad,$$

$$0\le m\le n+1\quad, \tag{A2.26}$$

and

$$E\,(x\,(n+1)-\hat{s}\,(n))\,i_x(m)\;=\;\gamma_m\,E\,(x\,(n+1)-\hat{s}\,(n))\,(x\,(m)-x\,(m-1))$$

$$=\;\gamma_m\,E\,(v\,(n+1)+w\,(n+1))\,(v\,(m)-v\,(m-1))\quad,$$

$$0\le m\le n\quad. \tag{A2.27}$$

Now the difference $x\,(n+1)-\hat{s}\,(n+1)$ is a linear combination of the innovations $i_x(m)$ for

$0 \le m \le n+1$.

$$x(n+1) - \hat{s}(n+1) = \sum_{m=0}^{n+1} h_m^{n+1} i_x(m) \tag{A2.28}$$

and the same is true for the difference $x(n+1) - \bar{s}(n)$,

$$x(n+1) - \bar{s}(n) = \sum_{m=0}^{n+1} g_m^{n+1} i_x(m) \quad . \tag{A2.29}$$

Combining (A2.26) — (A2.29) gives

$$h_m^{n+1} = \gamma_m E\, v(n+1)(v(m) - v(m-1)) \quad ,$$

$$0 \le m \le n+1 \quad , \tag{A2.30}$$

and

$$g_m^{n+1} = \gamma_m E\,(v(n+1) + w(n+1))(v(m) - v(m-1)) \quad ,$$

$$0 \le m \le n \quad . \tag{A2.31}$$

For the coefficient g_{n+1}^{n+1}, (A2.18) and (A2.29) give

$$g_{n+1}^{n+1} = \frac{1}{\gamma_{n+1}} \tag{A2.32}$$

since $\bar{s}(n)$ is a linear combination of $i_x(m)$ for $0 \le m \le n$ while $x(n+1)$ is a linear combination of $i_x(m)$ for $0 \le m \le n+1$.

Now

$$\frac{h_m^{n+1}}{g_m^{n+1}} = \begin{cases} \dfrac{E\, v(n+1)(v(m) - v(m-1))}{E\,(v(n+1) + w(n+1))(v(m) - v(m-1))} & , \quad 0 \le m \le n \quad , \\ \gamma_{n+1}^2 E\, v(n+1)(v(n+1) - v(n)) & , \quad m = n+1 \quad . \end{cases} \tag{A2.33}$$

Evaluation of the various expectations (from sections 2.4.3 and 3.6.1 — tedious details omitted) gives

$$E\,v(n+1)(v(m)-v(m-1)) = \frac{3}{4}\left(\frac{\beta_{n+2}^2}{4^{n+2-m}} + \cdots + \frac{\beta_Y^2}{4^{Y-m}} + \frac{\sigma_v^2}{4^{Y-m}}\right) \quad , \quad \text{(A2.34)}$$

$$E\,w(n+1)(v(m)-v(m-1)) = \frac{3}{4}\left(\frac{\beta_{n+1}^2}{4^{n+1-m}}\right) \quad , \quad \text{(A2.35)}$$

$$E\,v(n+1)(v(n+1)-v(n)) = \frac{3}{4}\left(\frac{\beta_{n+2}^2}{4} + \cdots + \frac{\beta_Y^2}{4^{Y-n-1}} + \frac{\sigma_v^2}{4^{Y-n-1}}\right) \quad , \quad \text{(A2.36)}$$

$$\gamma_{n+1}^2 = \frac{1}{F_{n+1}-F_n} = \frac{1}{\frac{3}{4}\left(\beta_{n+1}^2 + \frac{\beta_{n+2}^2}{4} + \cdots + \frac{\beta_Y^2}{4^{Y-n-1}} + \frac{\sigma_v^2}{4^{Y-n-1}}\right)} \quad . \quad \text{(A2.37)}$$

Substitution of these relations into (A2.33) yields

$$\frac{h_m^{n+1}}{g_m^{n+1}} = \frac{\frac{\beta_{n+2}^2}{4} + \cdots + \frac{\beta_Y^2}{4^{Y-n-1}} + \frac{\sigma_v^2}{4^{Y-n-1}}}{\beta_{n+1}^2 + \frac{\beta_{n+2}^2}{4} + \cdots + \frac{\beta_Y^2}{4^{Y-n-1}} + \frac{\sigma_v^2}{4^{Y-n-1}}} \quad ,$$

$$0 \le m \le n+1 \quad , \quad \text{(A2.38)}$$

or

$$\frac{h_m^{n+1}}{g_m^{n+1}} = e_{n+1} \quad ,$$

$$0 \le m \le n+1 \quad , \quad \text{(A2.39)}$$

which is independent of m.

Substituting into (A2.28) and (A2.29) gives

$$x(n+1) - \hat{s}(n+1) = \sum_{m=0}^{n+1} e_{n+1}\, g_m^{n+1}\, i_x(m) \qquad \text{(A2.40)}$$

and

$$x(n+1) - \hat{s}(n) = \sum_{m=0}^{n+1} g_m^{n+1}\, i_x(m) \quad , \qquad \text{(A2.41)}$$

or

$$x(n+1) - \hat{s}(n+1) = e_{n+1}(x(n+1) - \hat{s}(n)) \qquad \text{(A2.42)}$$

which may be expressed in the form of (3.77) as

$$\hat{s}(n+1) = e_{n+1}\hat{s}(n) + (1 - e_{n+1})\, x(n+1) \qquad \text{(A2.43)}$$

with the estimator coefficient e_{n+1} given by

$$e_{n+1} = \frac{\dfrac{\beta_{n+2}^2}{4} + \cdots + \dfrac{\beta_F^2}{4^{F-n-1}} + \dfrac{\sigma_v^2}{4^{F-n-1}}}{\beta_{n+1}^2 + \dfrac{\beta_{n+2}^2}{4} + \cdots + \dfrac{\beta_F^2}{4^{F-n-1}} + \dfrac{\sigma_v^2}{4^{F-n-1}}} \qquad \text{(A2.44)}$$

as in (3.78).

A2.3 Cholesky Factorisation of the Data Correlation Matrix and its Inverse[110]

With the Kalman whitening filter of the data given by (A2.17), it is easily shown that the inverse of the data correlation matrix, as given by (A1.9) in Appendix 1, is the product

$$R_l^{-1} = \Gamma_l^T \Gamma_l \tag{A2.45}$$

of the Kalman whitening filter and its matrix transpose.

Inversion of the whitening filter (easily performed by recursion on l) gives the Kalman innovations filter as

$$L_l = \Gamma_l^{-1} = \begin{bmatrix} \sqrt{F_0^-} & 0 & \cdots & 0 & 0 \\ \sqrt{F_0^-} & \sqrt{F_1 - F_0} & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sqrt{F_0^-} & \sqrt{F_1 - F_0} & \cdots & \sqrt{F_{l-1} - F_{l-2}} & 0 \\ \sqrt{F_0^-} & \sqrt{F_1 - F_0} & \cdots & \sqrt{F_{l-1} - F_{l-2}} & \sqrt{F_l - F_{l-1}} \end{bmatrix}, \tag{A2.46}$$

and it may be verified that the data correlation matrix R_l of (A1.3) is given as the product

$$R_l = L_l L_l^T \tag{A2.47}$$

of the innovations filter and its matrix transpose.

Note that both the whitening and innovations filters are lower-triangular; this is a consequence of the causality of both the Gram-Schmidt orthonormalisation, corresponding to the whitening filter, and of its inverse transformation, which corresponds to the innovations filter.

The decomposition of the correlation matrix and of its inverse into lower- and upper-triangular factors is known as ***Cholesky factorisation*** [110] and is the discrete equivalent of the spectral factorisation required in the design of Kalman filters for continuous signals. The lower-triangular factor represents a causal component (the whitening or innovations filters) while the upper-triangular factor corresponds to an anticausal component.

The continuous case involves the decomposition of the data spectral density function into left-half-plane (causal) and right-half-plane (anticausal) factors in the Laplace transform domain[110].

APPENDIX 3

PROBABILITY DENSITY FUNCTION OF THE ESTIMATOR COEFFICIENT

This derivation is based on that of the F density which appears in [8].

The chi-square (χ^2) density[8] is characteristic of a sum of squares of independent gaussian random variables. If

$$\chi^2 = y_1^2 + \cdots + y_M^2 \tag{A.1}$$

where the y_k are independent, identically-distributed (i.i.d.) gaussian variables with variance

$$E y_k^2 = \sigma_y^2 \tag{A.2}$$

then the variable χ^2 of (A.1) has probability density function (pdf) given by[8]

$$f_{\chi^2}(x) = \frac{1}{(2\sigma_y^2)^{M/2}\,\Gamma(\frac{M}{2})}\, x^{(\frac{M}{2}-1)}\, e^{\frac{-x}{2\sigma_y^2}}\, U(x) \tag{A.3}$$

which is a chi-square density with M degrees of freedom and variance parameter σ_y^2. $\Gamma(.)$ is the gamma, or generalised factorial, function[8].

Comparison with section 3.9.1 indicates that the variate $\hat{N}_{l+1}$, which is the *average* of 4^{l+1} squares of nominally i.i.d gaussian variables of variance $\sigma^2_{m_{l+1}}$, has the chi-square density

$$f_{\hat{N}_{l+1}}(x) = Mf_{\chi^2}(Mx) \tag{A.4}$$

or

$$f_{\hat{N}_{l+1}}(x) = \frac{M^{M/2}}{(2\sigma^2_{m_{l+1}})^{M/2}\Gamma(\frac{M}{2})} x^{(\frac{M}{2}-1)} e^{\frac{-Mx}{2\sigma^2_{m_{l+1}}}} U(x) \tag{A.5}$$

with $M = 4^{l+1}$ degrees of freedom and variance parameter $\sigma^2_{m_{l+1}}$.

Similarly, $\hat{N}_l$ has the chi-square density

$$f_{\hat{N}_l}(x) = \frac{P^{P/2}}{(2\sigma^2_{m_l})^{P/2}\Gamma(\frac{P}{2})} x^{(\frac{P}{2}-1)} e^{\frac{-Px}{2\sigma^2_{m_l}}} U(x) \tag{A.6}$$

with $P = 4^l$ degrees of freedom and variance parameter $\sigma^2_{m_l}$.

Assuming independence of $\hat{N}_{l+1}$ and $\hat{N}_l$, the density of the estimator coefficient

$$\hat{e}_l = \frac{\hat{N}_{l+1}}{4\hat{N}_l} \tag{A.7}$$

may be derived as follows.

Letting X and Y represent $\hat{N}_{l+1}$ and $\hat{N}_l$ respectively, then $\hat{e}_l = X/(4Y) = Z/4$. From the independence assumption, the joint density of X and Y is given by

$$f_{XY}(x,y) = f_X(x)f_Y(y) \tag{A.8}$$

where $f_X(x)$ and $f_Y(y)$ are as given by (A.5) and (A.6) respectively. Denoting the constant terms in these densities by k_1 and k_2,

$$f_{XY}(x,y) = k_1 k_2 x^{(\frac{M}{2}-1)} y^{(\frac{P}{2}-1)} e^{-(\frac{Mx}{2\sigma^2_{m_{i+1}}} + \frac{Py}{2\sigma^2_{m_i}})} U(x)\,U(y) \quad . \qquad \text{(A.9)}$$

A change of variable, replacing x by zy, gives

$$f_{XY}(zy,y) = k_1 k_2 z^{(\frac{M}{2}-1)} y^{(\frac{M+P}{2}-2)} e^{-(\frac{Mzy}{2\sigma^2_{m_{i+1}}} + \frac{Py}{2\sigma^2_{m_i}})} U(z)\,U(y) \quad . \qquad \text{(A.10)}$$

The marginal density of Z is then given by

$$f_Z(z) = \int_{-\infty}^{\infty} |y|\, f_{XY}(zy,y)\, dy \qquad \text{(A.11)}$$

or

$$f_Z(z) = k_1 k_2 z^{(\frac{M}{2}-1)} \int_0^{\infty} y^{(\frac{M+P}{2}-1)} e^{-(\frac{Mz}{2\sigma^2_{m_{i+1}}} + \frac{P}{2\sigma^2_{m_i}})y}\, dy \quad . \qquad \text{(A.12)}$$

Now from the definition of the gamma function,

$$\int_0^{\infty} x^b e^{-cx} dx = \frac{\Gamma(b+1)}{c^{b+1}} \qquad \text{(A.13)}$$

and so the integral in (A.12) reduces to

$$I = \frac{\Gamma(\frac{M+P}{2})}{\left[\frac{Mz}{2\sigma^2_{m_{i+1}}} + \frac{P}{2\sigma^2_{m_i}}\right]^{(\frac{M+P}{2})}} \qquad \text{(A.14)}$$

giving for the marginal density of Z

$$f_Z(z) = \frac{k_1 k_2 z^{(\frac{M}{2}-1)} \Gamma(\frac{M+P}{2})}{P^{(\frac{M+P}{2})} \left[\frac{1}{2\sigma_{m_i}^2} + \frac{(\frac{M}{P})z}{2\sigma_{m_{i+1}}^2}\right]^{(\frac{M+P}{2})}} . \tag{A.15}$$

Inserting the values of the constants k_1 and k_2 gives

$$f_Z(z) = \left[\frac{(\frac{M}{P})^{M/2} \Gamma(\frac{M+P}{2})}{\Gamma(\frac{M}{2}) \Gamma(\frac{P}{2})}\right] \left[\frac{1}{(2\sigma_{m_{i+1}}^2)^{\frac{M}{2}} (2\sigma_{m_i}^2)^{\frac{P}{2}}}\right] \left[\frac{z^{(\frac{M}{2}-1)}}{\left[\frac{1}{2\sigma_{m_i}^2} + \frac{(\frac{M}{P})z}{2\sigma_{m_{i+1}}^2}\right]^{(\frac{M+P}{2})}}\right] . \tag{A.16}$$

Now $\hat{e}_i = Z/4$ and so

$$f_{\hat{e}_i}(x) = 4 f_Z(4x) \tag{A.17}$$

or

$$f_{\hat{e}_i}(x) = \left[\frac{(\frac{4M}{P})^{M/2} \Gamma(\frac{M+P}{2})}{\Gamma(\frac{M}{2}) \Gamma(\frac{P}{2})}\right] \left[\frac{1}{(2\sigma_{m_{i+1}}^2)^{\frac{M}{2}} (2\sigma_{m_i}^2)^{\frac{P}{2}}}\right] \left[\frac{x^{(\frac{M}{2}-1)}}{\left[\frac{1}{2\sigma_{m_i}^2} + \frac{(\frac{4M}{P})x}{2\sigma_{m_{i+1}}^2}\right]^{(\frac{M+P}{2})}}\right] \tag{A.18}$$

which completes the derivation.

Quad-Tree Image Estimation: A New Image Model and its Application to Minimum Mean Square Error Image Restoration

S.C.Clippingdale, R.G.Wilson

University of Warwick, Coventry, U.K.

1. INTRODUCTION

Linear minimum mean square error (MMSE) methods have been used in signal estimation since the pioneering work of Wiener. More recently, state-space approaches based on Kalman filtering have been applied to a wide variety of filtering and prediction problems.

The application of such "optimal" methods to image enhancement has met with limited success, due to the inherently nonstationary character of images. The most successful methods reported to date involve a "multiple-model" approach to modeling an image, giving rise to a filter which adapts to local image structure[1]-[3]. While such modifications of classical methods have proven more successful than their simpler ancestors, it remains true to say that they are still largely based on adaptations of techniques developed for one-dimensional signals.

The methods described here differ in that they are based on a radically different statistical model of images. The use of quad-tree methods in image processing, although a comparatively recent development[4], has found increasing acceptance in applications ranging from data compression to segmentation[5]-[8]. However, to date most of these methods have been derived in a somewhat *ad hoc* fashion. It will be shown in this paper that such need not be the case.

In section 2 we describe the quad-tree structure and introduce a new class of image model which is defined on the tree. In the terms of this model, and for corruption of the image by additive white noise, we derive the optimal MMSE estimator of the image. It is shown that the estimator may be expressed as a recursion which is computationally highly efficient, achieving a reduction in computation greater than an order of magnitude over even a fairly small (7 x 7) spatial estimator.

Section 3 is concerned with the implementation of the method and its adaptation to "subjective" criteria of image "quality". The estimation of the model parameters - which determine the MMSE estimator coefficients - from noisy data is considered. We introduce a computationally inexpensive method of reducing the aliasing ("blocking") distortion which is an artifact of the quad-tree structure. Additionally, a modification which reduces the blurring of salient image features such as edges is demonstrated and results presented.

2. QUAD-TREE IMAGE MODEL AND ESTIMATOR

A. QUAD-TREE STRUCTURE

The quad-tree data structure is an example of a hierarchical or pyramidal organization, in which data is represented on a number of different levels. A given level l consists of a square array of nodes, and the level above it (level $l+1$) contains one node (the father) for every four (its children) on level l. Hence if level l is of dimension N x N, level $l+1$ is of dimension (N/2) x (N/2). The top level of the pyramid contains just one node, known as the root. Often the nodes on the lowest level (the leaves) are assigned the values of the pixels of an image to be processed in some way, and a father node is simply set equal to the average of its four children.

$$\bar{s}_{i+1,\,i,\,j} = \frac{1}{4}(\bar{s}_{i,\,2i,\,2j} + \bar{s}_{i,\,2i,\,2j+1} + \bar{s}_{i,\,2i+1,\,2j} + \bar{s}_{i,\,2i+1,\,2j+1}) \tag{1}$$

where the overbars indicate averages. The root node finally contains the average of all the image pixels. Defining the image (leaf) plane as level 0 and using an image of 256 x 256 pixels, the root node s_r is then on level 8 ($r = 8$).

B. MODEL DEFINITION

The model defines what is essentially the reverse process, that is the generation of an image from a process which operates down the tree, starting at the root. We use as the core of the model a rather general process, in which the value of a node is given as the value of its father plus a term which corresponds to a sample of a white noise of variance β_i^2 dependent upon level:

$$s_{i,\,2i+ii,\,2j+jj} = s_{i+1,\,i,\,j} + \beta_i w_{i,\,2i+ii,\,2j+jj}$$

$$ii\,,\,jj = 0\,,\,1$$

$$E\,w^2 = 1 \tag{2}$$

In order to generate an image with a mean value of zero, we require $s_r = 0$. The correlation between nodes is nonstationary,

$$E s_{i_a,\,i,\,j} s_{i_b,\,k,\,m} = E s_a^2 \tag{3}$$

where s_a is the lowest common ancestor of the two nodes.

We incorporate corruption by additive white gaussian noise of the image $s_{0\,i,\,j}$ (as generated by (2)) by setting the leaf node values to

$$\bar{s}_{0,\,i,\,j} = s_{0,\,i,\,j} + n_{i,\,j} \tag{4}$$

where $n_{i,\,j}$ is a sample of a white gaussian noise process with variance σ_n^2 which is uncorrelated with the image:

$$E n_{i,\,j} n_{k,\,m} = \sigma_n^2\,\delta_{i,\,k}\,\delta_{j,\,m} \tag{5}$$

$$E n_{i,\,j} s_{i,\,k,\,m} = 0 \tag{6}$$

Henceforth and throughout most of the paper, we drop the position indices so that $s_{i,\,j}$ becomes s_i . The context should avoid confusion.

Now from (2),

$$E s_0^2 = E s_r^2 + \beta_{r-1}^2 + \cdots + \beta_0^2 \tag{7}$$

and from (4)

$$E \bar{s}_0^2 = E s_0^2 + \beta_{r-1}^2 + \cdots + \beta_0^2 + \sigma_n^2$$
$$= E s_0^2 + \sigma_n^2 \tag{8}$$

We then form the tree $\bar{\mathbf{S}}$ of averages $\bar{s}$ as indicated by (1). The average energy of a node on level l is then

$$E \bar{s}_l^2 = E s_l^2 + \frac{\beta_{l-1}^2}{4} + \frac{\beta_{l-2}^2}{16} + \cdots + \frac{\beta_0^2}{4^l} + \frac{\sigma_n^2}{4^l} \tag{9}$$

C. THE ESTIMATOR

We now consider an estimator $\hat{s}_l$ of s_l . The MMSE criterion requires that $\hat{s}_l$ be a linear combination of data, and the coefficients of that combination are then given by the orthogonality principle[9].

The available data is the set of all nodes in the tree $\bar{\mathbf{S}}$. However, we wish to avoid spatial operations across a particular level and concentrate on "vertical" operations, i.e. those which are defined along branches in the tree only. (Without this constraint, the optimal estimator of s_0 would be a spatial, nonstationary -- and slow -- filter operating on level 0).

The implication is that we should estimate a node in terms only of itself and its ancestors in the tree:

$$\hat{s}_l = \sum_{k=l}^{r} \delta_k^l\,\bar{s}_k \tag{10}$$

The orthogonality principle gives the MMSE solution as

$$E(\hat{s}_l - s_l)\,\bar{s}_k = 0 \qquad (l \le k \le r) \tag{11}$$

or

$$E \hat{s}_l \bar{s}_k = E s_l s_k \qquad (l \le k \le r)\ . \tag{12}$$

The equations (12) may be expanded to give the full set of Yule - Walker equations for the estimator:

$$\delta_l^l \overline{Es_l^2} + \delta_{l-1}^l \overline{Es_{l-1}s_{l-1}} + \cdots + \delta_{r-1}^l \overline{Es_{r-1}s_{l+1}} + \delta_r^l \overline{Es_r s_l} = \overline{Es_l s_l}$$

$$\delta_l^l \overline{Es_{l-1}s_l} + \delta_{l-1}^l \overline{Es_{l-1}^2} + \cdots + \delta_{r-1}^l \overline{Es_{r-1}s_l} + \delta_r^l \overline{Es_{l-1}s_r} = \overline{Es_{l-1}s_l}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\delta_l^l \overline{Es_{l+1}s_r} + \delta_{l-1}^l \overline{Es_{l+1}s_{l-1}} + \cdots + \delta_{r-1}^l \overline{Es_{l+1}s_l} + \delta_r^l \overline{Es_{l+1}s_r} = \overline{Es_{l+1}s_l}$$

$$\delta_l^l \overline{Es_r s_r} + \delta_{l-1}^l \overline{Es_r s_{r-1}} + \cdots + \delta_{r-1}^l \overline{Es_r s_{l+1}} + \delta_r^l \overline{Es_r^2} = \overline{Es_r s_l} \qquad (13)$$

which may be expressed in matrix form as

$$\begin{vmatrix} \overline{Es_l^2} & \overline{Es_{l-1}s_l} & \cdots & \overline{Es_{r-1}s_l} & \overline{Es_r s_l} \\ \overline{Es_l s_{l-1}} & \overline{Es_{l-1}^2} & \cdots & \overline{Es_{r-1}s_{l-1}} & \overline{Es_r s_{l-1}} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \overline{Es_l s_{r-1}} & \overline{Es_{l-1}s_{r-1}} & \cdots & \overline{Es_{r-1}^2} & \overline{Es_r s_{r-1}} \\ \overline{Es_l s_r} & \overline{Es_{l-1}s_r} & \cdots & \overline{Es_{r-1}s_r} & \overline{Es_r^2} \end{vmatrix} \begin{vmatrix} \delta_l^l \\ \delta_{l-1}^l \\ \cdots \\ \delta_{r-1}^l \\ \delta_r^l \end{vmatrix} = \begin{vmatrix} \overline{Es_l s_l} \\ \overline{Es_{l-1}s_l} \\ \cdots \\ \overline{Es_{r-1}s_l} \\ \overline{Es_r s_l} \end{vmatrix} \qquad (14)$$

If we define the correlation matrix above as $\mathbf{A}_l$ and the vector of coefficients as $\vec{\Delta}_l^l$ then $\vec{\Delta}_l^l$ is given by

$$\vec{\Delta}_l^l = \mathbf{A}_l^{-1} \begin{vmatrix} \overline{Es_l s_l} \\ \overline{Es_{r-1}s_l} \\ \cdots \\ \overline{Es_{l+1}s_l} \\ \overline{Es_l s_l} \end{vmatrix} = \mathbf{A}_l^{-1} \vec{\nu} \qquad (15)$$

and the problem focuses on the inversion of $\mathbf{A}_l$. It follows from (1), (2) and (9) that we may express $\overline{Es_k s_m}$ ($k \geq m$ and remembering that $\overline{s}_k$ is an ancestor of $\overline{s}_m$) as

$$\overline{Es_k s_m} = \overline{Es_k^2} + \frac{\beta_{k-1}^2}{4} + \frac{\beta_{k-2}^2}{16} + \cdots + \frac{\beta_m^2}{4^{k-m}} + \frac{\sigma_n^2}{4^k}$$
$$= \overline{Es_k^2} \qquad (16)$$

So

$$\mathbf{A}_l = \begin{vmatrix} a_r & a_r & a_r & \cdots & a_r & a_r \\ a_r & a_{r-1} & a_{r-1} & \cdots & a_{r-1} & a_{r-1} \\ a_r & a_{r-1} & a_{r-2} & \cdots & a_{r-2} & a_{r-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_r & a_{r-1} & a_{r-2} & \cdots & a_{l+1} & a_{l+1} \\ a_r & a_{r-1} & a_{r-2} & \cdots & a_{l+1} & a_l \end{vmatrix} \qquad (17)$$

where $a_k = \overline{Es_k^2}$. The structure of this matrix is such that its inverse is tridiagonal:

$$\mathbf{A}_l^{-1} = \begin{vmatrix} \frac{a_{r-1}}{a_r(a_{r-1}-a_r)} & \frac{-1}{a_{r-1}-a_r} & \cdots & 0 & 0 \\ \frac{-1}{a_{r-1}-a_r} & \frac{a_{r-1}-a_r}{(a_{r-1}-a_{r-2})(a_{r-2}-a_r)} & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \frac{a_l-a_{l+2}}{(a_l-a_{l+1})(a_{l+1}-a_{l+2})} & \frac{-1}{a_l-a_{l+1}} \\ 0 & 0 & \cdots & \frac{-1}{a_l-a_{l+1}} & \frac{1}{a_l-a_{l+1}} \end{vmatrix} \qquad (18)$$

The vector $\vec{\nu}$ on the right of (15) may be expressed as

$$\vec{\nu} = \begin{vmatrix} \overline{Es_l s_l} \\ \overline{Es_{r-1}s_l} \\ \cdots \\ \overline{Es_{l+1}s_l} \\ \overline{Es_l s_l} \end{vmatrix} = \begin{vmatrix} a_r - b \\ a_{r-1} - 4b \\ \cdots \\ a_{l+1} - 4^{r-l-1}b \\ a_l - 4^{r-l}b \end{vmatrix} \qquad (19)$$

where

$$a_k = \overline{Es_k^2}$$
$$b = \frac{\beta_{l-1}^2}{4^{r-l+1}} + \frac{\beta_{l-2}^2}{4^{r-l+2}} + \cdots + \frac{\beta_0^2}{4^r} + \frac{\sigma_n^2}{4^r} \qquad (20)$$

Evaluation of (15) for a central ($l+1 \leq k \leq r-1$) element of $\vec{\Delta}_l^l$ gives, after extensive manipulation,

$$\delta_k^l = \frac{4^k \beta_k^2 Z_{l-1}}{Z_k Z_{k-1}} \qquad (l+1 \leq k \leq r-1) \quad (21)$$

where

$$Z_m = 4^m \beta_m^2 + 4^{m-1}\beta_{m-1}^2 + \cdots + \beta_0^2 + \sigma_n^2 \qquad (22)$$

Evaluation of (15) for the extreme elements δ_l^l and δ_r^l gives

$$\delta_l^l = \frac{4^l \beta_l^2}{Z_l} \qquad (23)$$

$$\delta_r^l = \frac{4^r \overline{Es_r^2}}{Z_r Z_{r-1}} Z_{l-1} \qquad (24)$$

which is consistent with (21) if we define

$$\beta_r^2 = E s_r^2 \quad . \tag{25}$$

We now have the solution vector $\vec{\Delta}_l$ for the optimal MMSE estimator $\bar{s}_l$ of s_l $(0 \le l \le r-1)$.

D. RECURSION

We wish to ascertain whether the estimator of (10) may be expressed in the recursive form

$$\bar{s}_l = \varepsilon_l^l \bar{s}_l + \varepsilon_{l+1}^l \hat{s}_{l+1} \tag{26}$$

with initial condition $\bar{s}_r = \bar{s}_r$. Iteration of (26) yields

$$\bar{s}_l = \sum_{k=l}^{r} \bar{s}_k \varepsilon_k^k \prod_{m=l}^{k-1} \varepsilon_{m+1}^m \tag{27}$$

which is equivalent to (10) if and only if

$$\varepsilon_k^k \prod_{m=l}^{k-1} \varepsilon_{m+1}^m = \delta_k^l \qquad (0 \le l \le r \ , \ l \le k \le r) \quad . \tag{28}$$

This condition may be expressed as

$$\varepsilon_l^l = \delta_l^l$$

$$\varepsilon_{l+1}^l = \frac{\delta_k^l}{\delta_k^{l+1}} \qquad (0 \le l \le r-1 \ , \ l+1 \le k \le r) \quad . \tag{29}$$

From (21) it is evident that

$$\frac{\delta_k^l}{\delta_k^{l+1}} = \frac{Z_{l-1}}{Z_l} \tag{30}$$

independent of k, and hence the condition is satisfied with

$$\varepsilon_{l+1}^l = \frac{Z_{l-1}}{Z_l} \tag{31}$$

with Z_l as given by (22), and

$$\varepsilon_l^l = \delta_l^l = \frac{4^l \beta_l^2}{Z_l} = 1 - \varepsilon_{l+1}^l \quad . \tag{32}$$

Hence the recursive solution is given finally by

$$\bar{s}_l = e_l \hat{s}_{l+1} + (1 - e_l)\bar{s}_l \tag{33}$$

where

$$e_l = d_{l+1} = \frac{4^{l-1}\beta_{l-1}^2 + 4^{l-2}\beta_{l-2}^2 + \cdots + 4\beta_l^2 + \beta_l^2 + \sigma_n^2}{4^l\beta_l^2 + 4^{l-1}\beta_{l-1}^2 + 4^{l-2}\beta_{l-2}^2 + \cdots + 4\beta_l^2 + \beta_l^2 + \sigma_n^2} \quad . \tag{34}$$

Note that

$$e_0 = \frac{\sigma_n^2}{\beta_0^2 + \sigma_n^2} \quad . \tag{35}$$

From (9) it is possible to simplify (34) to yield, for $1 \le l \le r-1$,

$$e_l = \frac{N_l}{4N_{l+1}} \tag{36}$$

where

$$N_k = \frac{3}{4}(\beta_{k-1}^2 + \frac{\beta_{k-2}^2}{4} + \cdots + \frac{\beta_0^2}{4^{k-1}} + \frac{\sigma_n^2}{4^{k-1}}) = E\bar{s}_{k-1}^2 - E\bar{s}_k^2 \quad . \tag{37}$$

If we define

$$N_0 = 3\sigma_n^2 \tag{38}$$

then (36) holds additionally for $l = 0$.

E. COMPUTATION

It is apparent from (33) that in calculating the estimate, two multiplications are required at each node. Since there are about $4N^2/3$ tree nodes for a N x N image, approximately 8/3 multiplications are required per image pixel. (The initial formation of the average-tree S involves only additions and trivial multiplications by 1/4, which may be accomplished by a double right-shift). This compares very favorably with, for example, even a small spatial filter. The authors of [2] use a filter mask of $15^2 = 225$ elements, requiring 225 multiplications per filter per pixel. The present scheme therefore achieves an acceleration approaching two orders of magnitude over such a filter, and almost one order of magnitude over even a (5 x 5) filter.

3. IMPLEMENTATION AND ADAPTATION TO VIEWING CRITERIA

A. PARAMETER ESTIMATION

It is apparent from (36) that in order to evaluate the estimator coefficients e_l we require the statistics N_l and N_{l+1} as given by (37). We do not, however, have access to these expected values and so must estimate them from the available data. This amounts to estimating the expected energies $E\bar{s}_l^2$ for each level l. The obvious estimator of $E\bar{s}_l^2$ is simply the sample energy

$$\hat{E}\bar{s}_l^2 = \underset{\text{level } l}{ave}(\bar{s}_l^2) \tag{39}$$

and the estimator of N_l is then

$$\hat{N}_l = \hat{E}\hat{s}_{l-1}^{-2} - \hat{E}\hat{s}_l^{-2} \quad . \tag{40}$$

Now (with position indices)

$$\frac{1}{4}\sum_{i=0}^{1}\sum_{j=0}^{1}\hat{s}_{l-1,\,2i+k,\,2j+l}^{-2} - \hat{s}_{l,\,i,\,j}^{-2} = \frac{1}{4}\sum_{k=0}^{1}\sum_{l=0}^{1} m\hat{s}_{l-1,\,2i+k,\,2j+l}^{-2} \tag{41}$$

where m is the difference

$$m_{k-1,\,2i+k,\,2j+l} - \bar{s}_{k-1,\,2i+k,\,2j+l} - \bar{s}_{k,\,i,\,j} \quad . \tag{42}$$

If we assume, as an approximation, that m_{k-1} has a gaussian density with zero mean and variance $\sigma^2_{m_{k-1}}$ then from (41) and [10], $\hat{N}_k$ has a chi-square density with $d_k = 4^{r-k+1}$ degrees of freedom and variance parameter $\sigma^2_{m_{k-1}}$. A similar statement may be made regarding $\hat{N}_{k+1}$. Then $\epsilon_k = \hat{N}_k / 4\hat{N}_{k+1}$ has the F density which is characteristic of a ratio of two chi-square variables[10]. The variance of ϵ_l thus calculated increases with increasing l as the sample size d_l decreases. Fortunately the problem is well conditioned in that the ϵ_l on the higher levels, while prone to larger errors in estimation, have proportionately less effect on the image estimate as is evident from the recursion of (33). We therefore use this simple method of estimating ϵ_l from the data and tolerate any error.

B. REDUCTION OF ALIAS (BLOCK) DISTORTION

The structure of the quad-tree, together with the upward averaging process (1) and downward propagation of data by the estimator (33), gives rise to noticeable block edge effects. These have been dramatically reduced in the present scheme by the inexpensive expedient of inserting intermediate nodes into the tree structure. These nodes do not themselves form a tree, but straddle the boundaries of the blocks which are artifacts of the original tree. On level l in the tree ($1 \le l \le r-1$) there are $(2^{r-l})^2$ original nodes $s_{(l,\,\cdot,\,\cdot)}$, and we now insert $(2^{r-l}-1)^2$ intermediate nodes $p_{(l,\,\cdot,\,\cdot)}$ such that each lies centrally between four original nodes. The "children" of an intermediate node are taken to be the four nodes closest to its vertical projection on to the level below. This implies that its children have four different original fathers (these being the four original nodes $s_{(l,\,\cdot,\,\cdot)}$ between which we have inserted the new node). Hence the new node straddles a block edge on level $l-1$. When we build the tree $\hat{\mathbf{S}}$ we insert at the new node the average $\bar{p}_l$ of its children in the usual way. Note that if the tree originally contained D nodes, there are approximately $D/4$ new nodes.

When we come to the downward estimation, we modify the estimator of the original nodes as

$$\hat{s}_l = e_l\left(\frac{\hat{s}_{l+1} + \hat{p}_{l+1}}{2}\right) + (1-e_l)\bar{s}_l \tag{43}$$

and the new nodes are estimated as

$$\hat{p}_l = e_l\hat{s}_{l+1} + (1-e_l)\bar{p}_l \tag{44}$$

where $\bar{s}_{l+1}$ is an average of the estimates of the surrounding original nodes $\hat{s}_l$ (p_l does not have a father in the sense that s_l does, so we create one).

C. EDGE PRESERVATION

It is well known[11] that edges and other linear features are particularly important to the visual system, and it has been observed that an image in which edges are preserved appears subjectively to be of higher "quality" than one in which the edges have been blurred; this is often true even when the blurred image has a superior signal to noise ratio (SNR). This constitutes a major disadvantage of stationary, "Wiener-type" image estimators which tend to blur the image, giving an unsatisfactory subjective appearance. A related and complementary phenomenon is the so-called "masking effect"[1], which describes the reduction of noise visibility in the vicinity of edges.

A number of estimation schemes[2][12] which take account of these effects have been devised. They seek to preserve edges by reducing the degree of smoothing of the image in their vicinity, where corrupting noise is less visible in any case.

The present scheme is amenable to the incorporation of such a modification. We have chosen once again, in the interests of speed, to reject spatial edge-detection operators and to concentrate on "vertical" methods defined on the tree structure.

Position indices have been omitted from what follows. We define at each node in the average-tree $\bar{S}$ an activity index γ_l $(l < r)$ as

$$\gamma_l = \frac{(\bar{s}_l - \bar{s}_{l+1})^2 + (\bar{s}_l - \bar{p}_{l+1})^2}{\hat{N}_{l+1}} \qquad (45)$$

and a propagated activity index α_l $(l < r)$ as

$$\alpha_l = \alpha_{l+1}^{(1-a_l)} \gamma_l^{a_l} \qquad (46)$$

$$\alpha_r = 1$$

where c_l $(0 \le c_l \le 1)$ is a weighting coefficient for the "linear" geometric combination of (46) which depends on the signal to noise ratio of the input image. We choose to use propagation (and a geometric rather than an arithmetic combination) in deference to the principle (Marr[11], Eklundh et al[13], Marr & Hildreth[14]) that an edge persists over a number of different scales (i.e. levels) whereas activity due to noise does not. The propagated geometric combination is thus a weighted "AND" which is influenced by activity on the higher levels more for low c_l (low SNR) and less for high c_l (high SNR). The coefficient c_l estimates the expected proportion of the activity γ_l at a given level l which is due to the uncorrupted image, and increases with l as the noise is progressively averaged out:

$$c_l = 1 - \frac{3}{4} \frac{\sigma_n^2}{4^l \hat{N}_{l+1}} \qquad (47)$$

The index α_l is used to alter the value of the estimator coefficient e_l :

$$e_l \leftarrow (e_l)^{\alpha_l} \qquad (48)$$

and e_l is finally used as the estimator coefficient in (43) and (44). The output image is scaled to the same energy as the uncorrupted image.

D. RESULTS

Table 1 lists input and output signal to noise ratios obtained with the present method. Figures 1 to 6 show the corresponding images as indicated in the table. The original image is shown in figure 7.

Table1

INPUT		OUTPUT	
SNR dB	Figure	SNR db	Figure
-18		4.1	
-12		6.7	
-6		9.2	
0	1	12.1	2
6	3	14.9	4
12	5	17.6	6
18		20.0	

The estimator produces quite acceptable results, with a signal to noise ratio improvement of 12dB for a 0dB input. The rather blotchy appearance of the "smooth" regions of the image, particularly evident in figure 2, is due to the edge detection system being fooled by the high level of noise at middle levels in the average-tree $\bar{S}$ and reducing the estimator coefficient accordingly, thus passing the offending noise down unsmoothed to the lower levels. Where the edge detection system responds correctly to an edge, such as around the face and shoulder, the (deliberately) unsmoothed noise from the input image is visible. However, the affected region is a rather wide strip along the edge at low input SNR. This is partly due to the operation of the weighted geometric mean (WGM) in the edge detector, which at low SNR takes more account of the father's activity and less of the child's. This effectively represents a smoothing of the edge index (in a manner similar to the smoothing of the signal performed by the estimator proper) and hence the WGM renders the edge detection machinery less susceptible to triggering by noise but at the expense of this spreading effect when there is a genuine edge at low SNR. That this should be the case is unsurprising when one considers that the uncertainty principle sets a limit on the performance of such systems, but it may be that a better (vertical?) edge detector would improve the appearance of the results at low SNR.

At higher SNR (such as figures 5 / 6) the system is capable of resolving the genuine edges with little difficulty and the output is appropriately smoothed in the smooth regions of the original without blurring the edges.

4. CONCLUSIONS

A new image model, based on the quad-tree data structure, has been presented. The model gives rise to a minimum mean square error (optimal) estimator which was shown to be considerably faster (by one or two orders of magnitude) than a conventional spatial filter in terms of the number of multiplications required per image pixel.

A modification designed to overcome the aliasing (block) distortion problem was presented. The insertion of additional nodes effectively creates a second alias which substantially cancels the original without the need for filtering prior to subsampling. The number of additional nodes required amounts to only some 25 percent of the original number, and the vertical nature of the algorithm remains largely intact.

The estimator was modified to incorporate an edge detection system, the image being smoothed less by the estimator in regions of edge activity. This measure takes account of the behavior of the visual system, which is usually the final arbiter of image "quality".

The overall system was shown to be effective in restoring imagery corrupted by additive white noise, achieving a respectable improvement in signal to noise ratio.

Currently the algorithm is isotropic, and there might be some utility in the development of a tree-based anisotropic estimator since the masking effect is anisotropic[2] in that it is affected by the orientation of the nearby edge. The edge detection scheme is ripe for improvement, particularly in its handling of input with poor signal to noise ratio.

A fast predictor/estimator is of potential utility in the coding of images; in such a system the coder might quantize and encode the difference $\hat{s}_{i-1} - \hat{s}_i$ instead of the difference $\bar{s}_{i-1} - \bar{s}_i$ which would normally be used in a quad-tree coder[7].

The addition of the intermediate nodes offers the possibility of developing a modified model which incorporates them. However,the structure is non-recursive and this greatly complicates its analysis. It is questionable whether the improvement in performance would be significant.

Further work might examine the application of this type of processor to more general pyramidal image models and data structures. The type of structure used in [6] (in which a node has four upward links and sixteen downward) could possibly serve as the basis for a (nearly) recursive model which could be free from the blocking effects of the "single-parent" quad-tree.

REFERENCES

[1] J.F.Abramatic, L.M.Silverman, *Nonlinear Restoration of Noisy Images*, IEEE Trans. P.A.M.I. Vol. PAMI-4, 141-9 (Mar 1982).

[2] H.E.Knutsson, R.G.Wilson, G.H.Granlund, *Anisotropic Nonstationary Image Estimation and its Applications Part I - Restoration of Noisy Images*, IEEE Trans. Commun. Vol. COM-31, 388-97 (Mar 1983).

[3] J.W.Woods, C.H.Radewan, *Kalman Filtering in Two Dimensions*, IEEE Trans. Inf. Theory Vol. IT-23, 473-82 (Jul 1977).

[4] S.L.Tanimoto, T.Pavlidis, *A Hierarchical Data Structure for Picture Processing*, Comp. Graphics and Im. Proc. 4 (1975).

[5] P.J.Burt, E.H.Adelson, *The Laplacian Pyramid as a Compact Image Code*, IEEE Trans. Commun. Vol. COM-31, 532-40 (Apr 1983).

[6] P.J.Burt, T.Hong, A.Rosenfeld, *Segmentation and Estimation of Image Region Properties Through Cooperative Hierarchical Computation*, IEEE Trans. Sys. Man Cyb. Vol. SMC-11, 802-9 (1981).

[7] R.G.Wilson, *Quad-Tree Predictive Coding: A New Class of Image Data Compression Algorithms*, Proc. IEEE Conf. on A.S.S.P., San Diego, CA (Mar 1984).

[8] M.Spann, R.G.Wilson, *A Quad-Tree Approach to Image Segmentation Which Combines Statistical and Spatial Information*, Pattern Recognition Vol. 18 Nos. 3/4, 257-69 (1985).

[9] A.Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York, 1965.

[10] D.R.Barr, P.W.Zehna, *Probability: Modeling Uncertainty*, Addison Wesley 1983.

[11] D.Marr, *Vision*, Freeman 1982.

[12] G.L.Anderson, A.N.Netravali, *Image Restoration based on a Subjective Criterion*, IEEE Trans. Sys. Man Cyb. Vol. SMC-6 pp 845 - 853, Dec 1976.

[13] J-O.Eklundh, T.Elfving, S.Nyberg, *Edge Detection using the Marr-Hildreth Operator with Different Sizes*, Proc. ICPR, Munich 1982.

[14] D.Marr, E.Hildreth, *Theory of Edge Detection*, MIT AI Memo 518, 1979.

FIGURES 1 - 7

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

REFERENCES

[1] H.B.Aasnaes, T.Kailath, *An Innovations Approach to Least-Squares Estimation: Part VII: Some Applications of Vector ARMA Models*, IEEE Trans. Vol. AC-18, 601-8 (1973)

[2] J.F.Abramatic, L.M.Silverman, *Nonstationary Linear Restoration of Noisy Images*, IEEE Conf. Dec. Control, Ft. Lauderdale, FL, 92-9 (1979)

[3] J.F.Abramatic, L.M.Silverman, *Nonlinear Restoration of Noisy Images*, IEEE Trans. Vol. PAMI-4, 141-9 (1982)

[4] E.H.Adelson, E.Simoncelli, R.Hingorani, *Orthogonal Pyramid Transforms for Image Coding*, Proc. SPIE Vis. Commun. and Im. Proc., Cambridge, MA (1987)

[5] N.Ahmed, K.R.Rao, *Orthogonal Transformations for Digital Signal Processing*, Springer Verlag, NY (1975)

[6] G.L.Anderson, A.N.Netravali, *Image Restoration Based on a Subjective Criterion*, IEEE Trans. Vol. SMC-6, 845-53 (1976)

[7] T.Assefi, *Stochastic Processes and Estimation Theory with Applications*, Wiley (1979)

[8] D.R.Barr, P.W.Zehna, *Probability: Modeling Uncertainty*, Addison-Wesley (1983)

[9] C.Blakemore, F.W.Campbell, *On the Existence in the Human Visual System of Neurons Selectively Sensitive to the Orientation and Size of Retinal Images*, J. Physiol., Vol. 203, 237-60 (1969)

[10] Z.L.Budrikis, *Visual Fidelity Criterion and Modeling*, Proc. IEEE Vol. 60, 771-9 (1972)

[11] P.J.Burt, E.H.Adelson, *The Laplacian Pyramid as a Compact Image Code*, IEEE Trans. Vol. COM-31, 532-40 (1983)

[12] P.J.Burt, T.H.Hong, A.Rosenfeld, *Segmentation and Estimation of Image Region Properties through Cooperative Hierarchical Computation*, IEEE Trans. Vol. SMC-11, 802-9 (1981)

[13] J.C.Candy, R.H.Bosworth, *Methods for Designing Differential Quantizers based on Subjective Evaluations of Edge Busyness*, Bell Syst. Tech. J. Vol. 51, 1495-1516 (1972)

[14] J.F.Canny, *A Computational Approach to Edge Detection*, IEEE Trans. Vol. PAMI-8, 679-98 (1986)

[15] V.Cantoni, S.Levialdi, eds., *Pyramidal Systems for Computer Vision*, Springer Verlag, Berlin (1986)

[16] R.Chellappa, R.L.Kashyap, *Digital Image Restoration Using Spatial Interaction Models*, IEEE Trans. ASSP-30, 461-72 (1982)

[17] P.C.Chen, T.Pavlidis, *Image Segmentation as an Estimation Problem*, Comp. Graph. Im. Proc. Vol. 12, 153-72 (1980)

[18] J.C.Chow, *On Estimating the Orders of an ARMA Process with Uncertain Observations*, IEEE Trans. Vol. AC-17, 707-9 (1972)

[19] S.C.Clippingdale, R.G.Wilson, *Quad-Tree Image Estimation: A New Image Model and its Application to MMSE Image Restoration*, Proc. 5th Scandinavian Conf. Im. Analysis, Stockholm, 699-706 (1987)

[20] F.S.Cohen, D.B.Cooper, *Simple Parallel Hierarchical and Relaxation Algorithms for Segmenting Noncausal Markovian Random Fields*, IEEE Trans. Vol. PAMI-9, 195-219 (1987)

[21] Y.Cohen, M.S.Landy, M.Pavel, *Hierarchical Coding of Binary Images*, IEEE Trans. Vol. PAMI-7, 284-97 (1985)

[22] G.R.Cross, A.K.Jain, *Markov Random Field Texture Models*, IEEE Trans. Vol. PAMI-5, 25-39 (1983)

[23] J.G.Daugman, *Six Formal Properties of Two-Dimensional Anisotropic Visual Filters: Structural Principles and Frequency/Orientation Selectivity*, IEEE Trans. Vol. SMC-13, 882-7 (1983)

[24] J.G.Daugman, *Uncertainty Relation for Resolution in Space, Spatial Frequency and Orientation Optimised by Two-Dimensional Visual Cortical Filters*, J. Opt. Soc. Am. Vol. A2, 1160-9 (1985)

[25] R.Deutsch, *Estimation Theory*, Prentice Hall (1966)

[26] J-O.Eklundh, T.Elfving, S.Nyberg, *Edge Detection Using the Marr-Hildreth Operator with Different Sizes*, Proc. Int. Conf. Patt. Rec., München, 1109-12 (1982)

[27] M.P.Ekstrom, *Realisable Wiener Filtering in Two Dimensions*, IEEE Trans. Vol. ASSP-30, 31-40 (1982)

[28] M.P.Ekstrom, R.E.Twogood, J.W.Woods, *Two-Dimensional Recursive Filter Design: A Spectral Factorization Approach*, IEE... Trans. Vol. ASSP-28, 16-26 (1980)

[29] M.P.Ekstrom, J.W.Woods, *Two-Dimensional Spectral Factorization with Applications in Recursive Digital Filtering*, IEEE Trans. Vol. ASSP-24, 115-28 (1976)

[30] C.Enroth-Cugell, J.G.Robson, *The Contrast Sensitivity of Retinal Ganglion Cells of the Cat*, J. Physiol. Vol. 187, 517-52 (1966)

[31] K-B.Eom, R.L.Kashyap, *Texture and Intensity Edge Detection with Random Field Models*, Proc. IEEE Workshop on Computer Vision, Miami, FL, 29-34 (1987)

[32] J.R.Fram, E.S.Deutsch, *On the Quantitative Evaluation of Edge Detection Schemes and their Comparison with Human Performance*, IEEE Trans. Vol. C-24, 616-28 (1975)

[33] W.Frei, C.C.Chen, *Fast Boundary Detection: A Generalisation and a New Algorithm*, IEEE Trans. Vol. C-26, 988-98 (1977)

[34] N.C.Gallagher Jr., G.L.Wise, *A Theoretical Analysis of the Properties of Median Filters*, IEEE Trans. Vol. ASSP-29, 1136-41 (1981)

[35] W.Gersch, D.R.Sharpe, *Estimation of Power Spectra with Finite Order Autoregressive Models*, IEEE Trans. Vol. AC-18, 367-9 (1973)

[36] D.M.Goodman, M.P.Ekstrom, *Multidimensional Spectral Factorization and Unilateral Autoregressive Models*, IEEE Trans. Vol. AC-25, 258-62 (1980)

[37] A.Graham, *Kronecker Products and Matrix Calculus with Applications*, Wiley (1981)

[38] G.H.Granlund, *In Search of a General Picture Processing Operator*, Comp. Graph. Im. Proc. Vol. 8, 155-73 (1978)

[39] R.N.Haber, M.Hershenson, *The Psychology of Visual Perception*, Holt, Rinehart, Winston, NY (1980)

[40] A.Habibi, *Hybrid Coding of Pictorial Data*, IEEE Trans. Vol. COM-22, 614-24 (1974)

[41] A.Habibi, *A Survey of Adaptive Image Coding Techniques*, IEEE Trans. Vol. COM-25, 1275-84 (1977)

[42] A.Habibi, R.S.Hershel, *A Unified Representation of DPCM and Transform Coding Systems*, IEEE Trans. Vol. COM-22, 692-6 (1974)

[43] A.Habibi, *Two-Dimensional Bayesian Estimate of Images*, Proc. IEEE Vol. 60, 878-83 (1972)

[44] R.M.Haralick, K.S.Shanmugam, I.H.Dinstein, *Textural Features for Image Classification*, IEEE Trans. Vol. SMC-3, 610-21 (1973)

[45] L.Harmon, *The Recognition of Faces*, Sci. Am., Nov. 1973

[46] W.Heisenberg, *The Physical Principles of the Quantum Theory*, Dover, NY (1930)

[47] C.W.Helstrom, *Image Restoration by the Method of Least Squares*, J. Opt. Soc. Am. Vol. 57, 297-303 (1967)

[48] T.H.Hong, K.A.Narayanan, S.Peleg, A.Rosenfeld, T.Silberberg, *Image Smoothing and Segmentation by Multiresolution Pixel Linking: Further Experiments and Extensions*, IEEE Trans. Vol. SMC-12, 611-22 (1982)

[49] T.H.Hong, A.Rosenfeld, *Compact Region Extraction Using Weighted Pixel Linking in a Pyramid*, IEEE Trans. Vol. PAMI-6, 222-9 (1984)

[50] T.H.Hong, M.Shneier, *Extracting Compact Objects Using Linked Pyramids*, IEEE Trans. Vol. PAMI-6, 229-37 (1984)

[51] T.H.Hong, M.Shneier, A.Rosenfeld, *Border Extraction Using Linked Edge Pyramids*, IEEE Trans. Vol. SMC-12, 660-8 (1982)

[52] T.S.Huang, W.F.Schreiber, O.J.Tretiak, *Image Processing*, Proc. IEEE Vol. 59, 1586-1609 (1971)

[53] T.S.Huang, G.J.Yang, G.Y.Tang, *A Fast Median Filtering Algorithm*, IEEE Trans. Vol. ASSP-27, (1979)

[54] D.H.Hubel, T.N.Wiesel, *Receptive Fields of Single Neurones in the Cat's Striate Cortex*, J. Physiol. Vol. 148, 574-91 (1959)

[55] D.H.Hubel, T.N.Wiesel, *Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex*, J. Physiol. Vol. 160, 106-54 (1962)

[56] D.H.Hubel, T.N.Wiesel, *Brain Mechanisms of Vision*, Sci. Am., Sept. 1979

[57] B.R.Hunt, *The Application of Constrained Least Squares Estimation to Image Restoration by Digital Computer*, IEEE Trans. Vol. C-22, 805-12 (1973)

[58] G.M.Hunter, K.Steiglitz, *Operations on Images Using Quadtrees*, IEEE Trans. Vol.PAMI-1, 145-53 (1979)

[59] V.K.Ingle, J.W.Woods, *Multiple Model Recursive Estimation of Images*, Proc. ICASSP'79, Washington DC, 642-5 (1979)

[60] A.K.Jain, E.Angel, *Image Restoration, Modeling and Reduction of Dimensionality*, IEEE Trans. Vol. C-23, 470-6 (1974)

[61] A.K.Jain, *A Semicausal Model for Recursive Filtering of Two-Dimensional Images*, IEEE Trans. Vol. C-26, 343-50 (1977)

[62] A.K.Jain, *A Fast Karhunen-Loeve Transform for Recursive Filtering of Images Corrupted by White and Colored Noise*, IEEE Trans. Vol. C-26, 560-71 (1977)

[63] A.K.Jain, *Image Data Compression: A Review*, Proc. IEEE Vol. 69, 349-89 (1981)

[64] A.K.Jain, J.R.Jain, *Partial Differential Equations and Finite Difference Methods in Image Processing: Part II: Image Restoration*, IEEE Trans. Vol. AC-23, 817-34 (1978)

[65] A.K.Jain, *Advances in Mathematical Models for Image Processing*, Proc. IEEE Vol. 69, 502-28 (1981)

[66] A.K.Jain, *A Sinusoidal Family of Unitary Transforms*, IEEE Trans. Vol. PAMI-1, 356-65 (1979)

[67] T.Kailath, *An Innovations Approach to Least Squares Estimation: Part I: Linear Filtering in Additive White Noise*, IEEE Trans. Vol. AC-13, 646-55 (1968)

[68] T.Kailath, *The Innovations Approach to Detection and Estimation Theory*, Proc. IEEE Vol. 58, 680-95 (1970)

[69] T.Kailath, *A View of Three Decades of Linear Filtering Theory*, IEEE Trans. Vol. IT-20, 146-81 (1974)

[70] R.E.Kalman, R.S.Bucy, *New Results in Linear Filtering and Prediction Theory*, Trans. ASME J. Bas. Eng. Vol. D83, 95-107 (1961)

[71] R.E.Kalman, *A New Approach to Linear Filtering and Prediction Problems*, Trans. ASME J. Bas. Eng. Vol. D82, 35-45 (1960)

[72] R.L.Kashyap, *Univariate and Multivariate Random Field Models for Images*, Comp. Graph. Im. Proc. Vol. 12, 257-70 (1980)

[73] K.Knowlton, *Progressive Transmission of Grey-Scale and Binary Pictures by Simple, Efficient and Lossless Encoding Schemes*, Proc. IEEE Vol. 68, 885-96 (1980)

[74] D.E.Knuth, *The Art of Computer Programming, Vol. 2, Seminumerical Algorithms*, 2nd ed. Addison Wesley (1981)

[75] H.Knutsson, R.C.Wilson, G.H.Granlund, *Anisotropic Nonstationary Image Estimation and Its Applications: Part I - Restoration of Noisy Images*, IEEE Trans. Vol. COM-31, 388-97 (1983)

[76] R.G.Wilson, H.Knutsson, G.H.Granlund, *Anisotropic Nonstationary Image Estimation and Its Applications: Part II - Predictive Image Coding*, IEEE Trans. Vol. COM-31, 398-407 (1983)

[77] H.Knutsson, R.G.Wilson, G.H.Granlund, *Anisotropic Filtering Operations for Image Enhancement and their Relation to the Visual System*, Proc. IEEE Conf. Pat. Rec. Im. Proc., Dallas, TX (1981)

[78] H.Knutsson, R.G.Wilson, G.H.Granlund, *Estimating the Local Orientation of Anisotropic Two-Dimensional Signals*, Proc. IEEE ASSP Spectrum Estimation Workshop, Tampa, FL, 234-9 (1983)

[79] H.Knutsson, *Filtering and Reconstruction in Image Processing*, Linkoping University Press, Sweden (1982)

[80] H.Knutsson, *Filtering and Reconstruction in Image Processing*, Ph.D. dissertation, Linkoping University, Sweden (1982)

[81] J.J.Kulikowski, P.E.King-Smith, *Spatial Arrangement of Line, Edge and Grating Detectors Revealed by Subthreshold Summation*, Vision Res. Vol. 13, 1455-78 (1973)

[82] P.E.King-Smith, J.J.Kulikowski, *The Detection and Recognition of Two Lines*, Vision Res. Vol. 21, 235-50 (1981)

[83] D.S.Lebedev, L.I.Mirkin, *Smoothing of Two-Dimensional Images Using the "Composite" Model of a Fragment*, in *Iconics, Digital Holography and Image Processing*, Inst. for Problems in Information Transmission, Acad. of Sciences, USSR, 57-62 (1975)

[84] N.Levinson, *The Wiener RMS Error Criterion in Filter Design and Prediction*, J. Math. Phys. Vol. 25, 261-78 (1947)

[85] R.Linsker, *Self-Organisation in a Perceptual Network*, IEEE *Computer*, Vol. 21, No. 3, 105-17, March 1988

[86] J.Makhoul, *Linear Prediction: A Tutorial Review*, Proc. IEEE Vol. 63, 561-80 (1975)

[87] B.B.Mandelbrot, *The Fractal Geometry of Nature*, Freeman, San Francisco, CA (1982)

[88] J.L.Mannos, D.J.Sakrison, *The Effects of a Visual Fidelity Criterion on the Encoding of Images*, IEEE Trans. Vol. IT-20, 525-36 (1974)

[89] S.Marcelja, *Mathematical Description of the Responses of Cortical Simple Cells*, J. Opt. Soc. Am. Vol. 70, 1297-1300 (1980)

[90] D.Marr, E.Hildreth, *Theory of Edge Detection*, Proc. R. Soc. Lond. Vol. B207, 187-217 (1980)

[91] D.Marr, *Vision*, Freeman, San Francisco, CA (1982)

[92] P.Meer, E.S.Bauer, A.Rosenfeld, *Frequency Domain Analysis and Synthesis of Image Pyramid Generating Kernels*, IEEE Trans. Vol. PAMI-9, 512-22 (1987)

[93] J.L.Melsa, D.L.Cohn, *Decision and Estimation Theory*, McGraw-Hill (1978)

[94] H.Mostafavi, D.J.Sakrison, *Structure and Properties of a Single Channel in the Human Visual System*, Vision Res. Vol. 16, 957-68 (1976)

[95] J.A.Movshon, I.D.Thompson, D.J.Tolhurst, *Receptive Field Organisation of Complex Cells in the Cat's Striate Cortex*, J. Physiol. Vol. 283, 79-99 (1978)

[96] J.A.Movshon, I.D.Thompson, D.J.Tolhurst, *Spatial Summation in the Receptive Fields of Simple Cells in the Cat's Striate Cortex*, J. Physiol. Vol. 283, 53-77 (1978)

[97] M.S.Murphy, L.M.Silverman, *Image Model Representation and Line-by-Line Recursive Restoration*, IEEE Trans. Vol. AC-23, 809-16 (1978)

[98] N.E.Nahi, T.Assefi, *Bayesian Recursive Image Estimation*, IEEE Trans. Vol. C-21, 734-8 (1972)

[99] N.E.Nahi, *Estimation Theory and Applications*, Wiley (1969)

[100] N.E.Nahi, C.A.Franco, *Recursive Image Enhancement by Vector Scanning*, IEEE Trans. Vol. COM-21, 305-11 (1973)

[101] N.E.Nahi, A.Habibi, *Decision Directed Recursive Image Enhancement*, IEEE Trans. Vol. CAS-22, 286-93 (1975)

[102] N.E.Nahi, *The Role of Recursive Estimation in Statistical Image Enhancement*, Proc. IEEE Vol. 60, 872-77 (1972)

[103] H.Nakayama, M.Sone, M.Takagi, *Meteorological Satellite Image Analysis Using Fractal Dimension and Lower-Order Statistics*, Proc. 5th Scandinavian Conf. Im. Analysis, Stockholm, 261-8 (1987)

[104] P.M.Narendra, *A Separable Median Filter for Noise Smoothing*, IEEE Trans. Vol. PAMI-3, 20-9 (1981)

[105] A.N.Netravali, J.O.Limb, *Picture Coding: A Review*, Proc. IEEE Vol. 68, 366-406 (1980)

[106] J.A.Noble, *Finding Two-Dimensional Image Structure*, Proc. IEEE Workshop on Computer Vision, Miami, FL, 222-4 (1987)

[107] A.V.Oppenheim, R.W.Schafer, *Digital Signal Processing*, Prentice Hall (1975)

[108] A.V.Oppenheim, R.W.Schafer, T.G.Stockham Jr., *The Nonlinear Filtering of Multiplied and Convolved Signals*, Proc. IEEE Vol. 56, 1264-91 (1968)

[109] D.P.Panda, A.C.Kak, *Recursive Least Squares Smoothing of Noise in Images*, IEEE Trans. Vol. ASSP-25, 520-4 (1977)

[110] A.Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 2nd ed. (1984)

[111] A.Papoulis, *Signal Analysis*, McGraw-Hill (1981)

[112] S.Peleg, J.Naor, R.Hartley, D.Avnir, *Multiple Resolution Texture Analysis and Classification*, IEEE Trans. Vol. PAMI-6, 518-23 (1984)

[113] A.P.Pentland, *Fractal-Based Descriptions of Natural Scenes*, IEEE Trans. Vol. PAMI-6, 661-74 (1984)

[114] M.Pietikainen, A.Rosenfeld, *Image Segmentation by Texture Using Pyramid Node Linking*, IEEE Trans. Vol. SMC-11, 822-5 (1981)

[115] D.A.Pollen, S.F.Ronner, *Visual Cortical Neurons as Localised Spatial Frequency Filters*, IEEE Trans. Vol. SMC-13, 907-16 (1983)

[116] D.A.Pollen, S.F.Ronner, *Spatial Computation Performed by Simple and Complex Cells in the Visual Cortex of the Cat*, Vision Res. Vol. 22, 104-18 (1982)

[117] S.R.Powell, L.M.Silverman, *Modeling of Two-Dimensional Covariance Functions With Application to Image Restoration*, IEEE Trans. Vol. AC-19, 8-13 (1974)

[118] W.K.Pratt, *Generalised Wiener Filtering Computation Techniques*, IEEE Trans. Vol. C-21, 636-41 (1972)

[119] W.K.Pratt, O.D.Faugeras, A.Gagalowicz, *Applications of Stochastic Texture Field Models to Image Processing*, Proc. IEEE Vol. 61, 542-51 (1981)

[120] L.R.Rabiner, B.Gold, *Theory and Application of Digital Signal Processing*, Prentice Hall (1975)

[121] S.A.Rajala, R.J.P.DeFigueirdo, *Adaptive Nonlinear Image Restoration by a Modified Kalman Filtering Approach*, IEEE Trans. Vol. ASSP-29, 1033-42 (1981)

[122] S.Ranade, M.Shneier, *Using Quadtrees to Smooth Images*, IEEE Trans. Vol. SMC-11, 373-6 (1981)

[123] A.Rosenfeld, M.Thurston, *Edge and Curve Detection for Visual Scene Analysis*, IEEE Trans. Vol. C-20, 562-9 (1971)

[124] D.J.Sakrison, *On the Role of the Observer and a Distortion Measure in Image Transmission*, IEEE Trans. Vol. COM-25, 1251-66 (1977)

[125] B.G.Schunk, *Edge Detection With Gaussian Filters at Multiple Scales*, Proc. IEEE Workshop on Computer Vision, Miami, FL, 208-10 (1987)

[126] K.S.Shanmugam, F.M.Dickey, J.A.Green, *An Optimal Frequency-Domain Filter for Edge Detection in Digital Pictures*, IEEE Trans. Vol. PAMI-1, 37-69 (1979)

[127] F.C.Schoute, M.F.TerHorst, J.C.Willems, *Hierarchic Recursive Image Enhancement*, IEEE Trans. Vol. CAS-24, 67-78 (1977)

[128] D.Slepian, *Linear Least Squares Filtering of Distorted Images*, J. Opt. Soc. Am. Vol. 57, 297-303 (1967)

[129] M.M.Sondhi, *Image Restoration: The Removal of Spatially Invariant Degradations*, Proc. IEEE Vol. 60, 842-53 (1972)

[130] M.Spann, R.G.Wilson, *A Quadtree Approach to Image Segmentation Which Combines Statistical and Spatial Information*, Patt. Rec. Vol. 18, 257-69 (1985)

[131] M.Spann, *Texture Description and Segmentation in Image Processing*, Ph.D. dissertation, University of Aston (1985)

[132] T.G.Stockham Jr., *Image Processing in the Context of a Visual Model*, Proc. IEEE Vol. 60, 828-42 (1972)

[133] M.G.Strintzis, *Dynamic Representation and Recursive Estimation of Cyclic and Two-Dimensional Processes*, IEEE Trans. Vol. AC-23, 801-9 (1978)

[134] M.G.Strintzis, *On the Spatially Causal Estimation of Two-Dimensional Processes*, Proc. IEEE Vol. 65, 979-81 (1977)

[135] J.A.Stuller, B.Kurz, *Two-Dimensional Markov Representations of Sampled Images*, IEEE Trans. Vol. COM-24, 1148-52 (1976)

[136] S.L.Tanimoto, T.Pavlidis, *A Hierarchical Data Structure for Picture Processing*, Comp. Graph. Im. Proc. Vol. 4, 104-19 (1975)

[137] H.Taub, D.L.Schilling, *Principles of Communications Systems*, McGraw-Hill (1971)

[138] S.A.Tretter, K.Steiglitz, *Power Spectrum Identification in Terms of Rational Models*, IEEE Trans. Vol. AC-12, 185-88 (1967)

[139] *Unix Programmer's Manual*, UC Berkeley (1985)

[140] A.B.Watson, A.J.Ahumada Jr., *An Orthogonal Oriented Quadrature Hexagonal Image Pyramid*, NASA Tech. Memo. 100054 (1987)

[141] A.B.Watson, *The Cortex Transform: Rapid Computation of Simulated Neural Images*, Comp. Vision, Graph., Im. Proc. Vol. 39, 311-27 (1987)

[142] A.B.Watson, *Ideal Shrinking and Expansion of Discrete Sequences*, NASA Tech. Memo. 88202 (1986)

[143] N.Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series*, Wiley, NY (1949)

[144] A.S.Willsky, *Fourier Series and Estimation on the Circle*, IEEE Trans. Vol. IT-20, 577-83 (1974)

[145] R.G.Wilson, A.D.Calway, *A General Multiresolution Signal Descriptor and Its Application to Image Analysis*, Proc. Eusipco, Grenoble (1988)

[146] R.G.Wilson, G.H.Granlund, *The Uncertainty Principle in Image Processing*, IEEE Trans. Vol. PAMI-6, 758-67 (1984)

[147] R.G.Wilson, *A Class of Local Centroid Algorithms for Classification and Quantisation in Spaces of Arbitrary Dimension*, Linköping University, Tech. Rep. LiTH-ISY-I-0610 (1983)

[148] R.G.Wilson, *Quadtree Predictive Coding*, Proc. ICASSP'84, San Diego, CA (1984)

[149] R.G.Wilson, M.Spann, *Image Segmentation and the Problem of Uncertainty*, Research Studies Press (1987)

[150] R.G.Wilson, M.Spann, *Finite Prolate Spheroidal Sequences and their Applications: Part II: Image Feature Description and Segmentation*, IEEE Trans. Vol. PAMI-10, 193-203 (1988)

[151] A.P.Witkin, *Scale-Space Filtering*, Proc. ICASSP'84, San Diego, CA (1984)

[152] J.W.Woods, S.Dravida, R.Mediavilla, *Image Estimation Using Doubly Stochastic Gaussian Random Field Models*, IEEE Trans. Vol. PAMI-9, 245-53 (1987)

[153] J.W.Woods, *Two-Dimensional Discrete Markovian Fields*, IEEE Trans. Vol. IT-18, 232-40 (1972)

[154] J.W.Woods, *Two-Dimensional Markov Spectral Estimation*, IEEE Trans. Vol. IT-22, 552-9 (1976)

[155] J.W.Woods, V.K.Ingle, *Kalman Filtering in Two Dimensions: Further Results*, IEEE Trans. Vol. ASSP-29, 188-97 (1981)

[156] J.W.Woods, C.H.Radewan, *Kalman Filtering in Two Dimensions*, IEEE Trans. Vol. IT-23, 473-82 (1977)

[157] J.W.Woods, *Two-Dimensional Kalman Filtering*, in *Two-Dimensional Transforms and Filters*, T.S.Huang, ed., Springer Verlag, Berlin (1981)

[158] J.K.Yan, D.J.Sakrison, *Encoding of Images Based on a Two-Component Source Model*, IEEE Trans. Vol. COM-25, 1315-22 (1977)

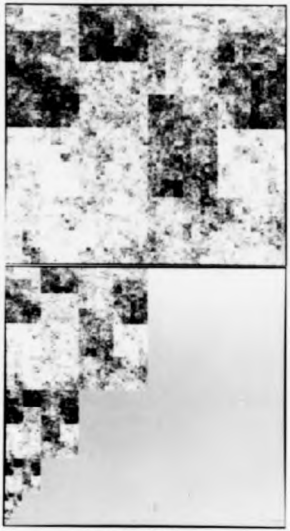


Figure P1

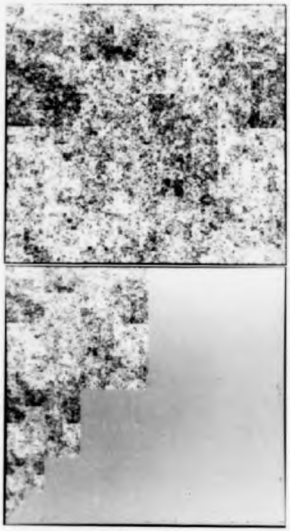


Figure P2

Figure P3

Figure P4

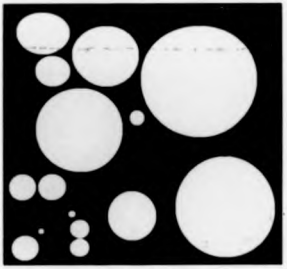


Figure P5

Figure P6

Figure P7

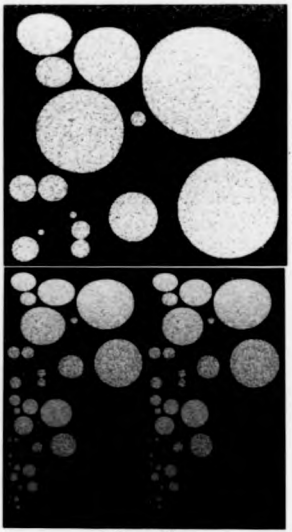


Figure P8

Figure P9

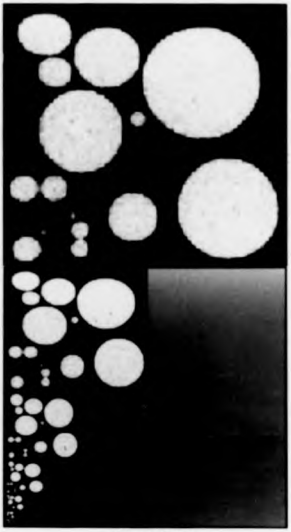


Figure P10

Figure P11

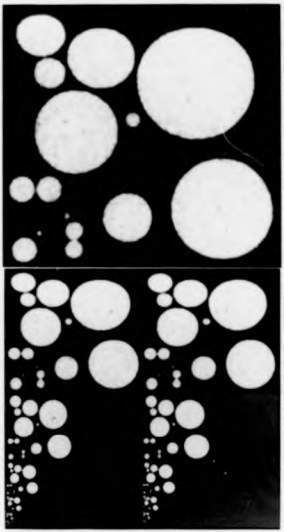


Figure P12

Figure P13

Figure P14

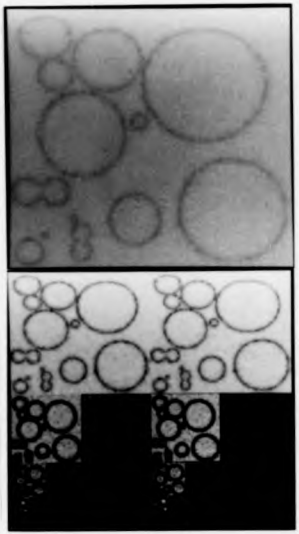


Figure P15

Figure P16

Figure P17

Figure P18

Figure P19

Figure P20

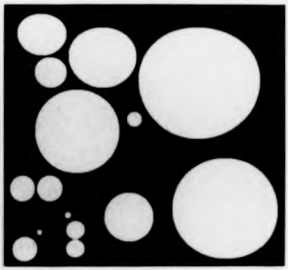


Figure P21

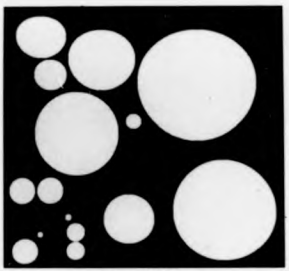


Figure P22

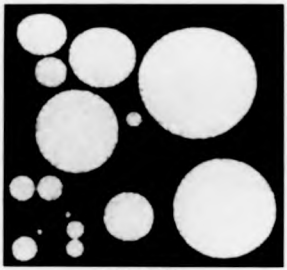


Figure P23

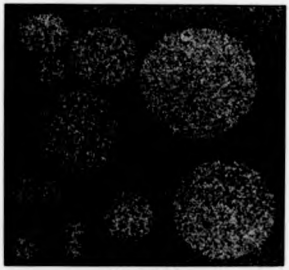


Figure P24

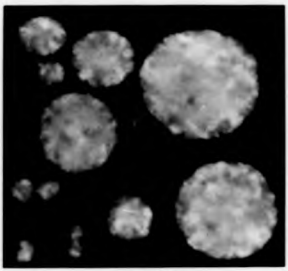


Figure P25

Figure P26

Figure P27

Figure P28

Figure P29

Figure P30

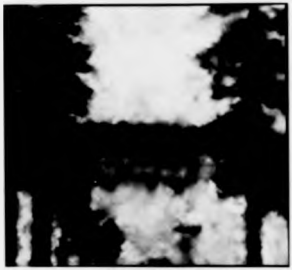


Figure P31

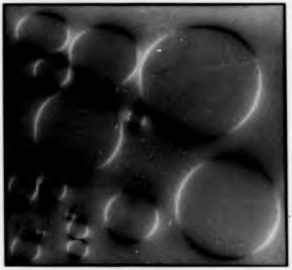


Figure P34

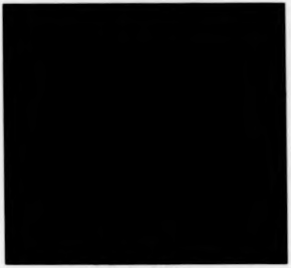


Figure P32

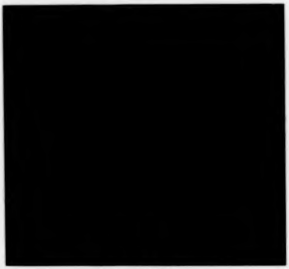


Figure P35

Figure P33

Figure P36

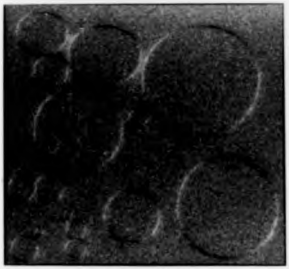


Figure P37

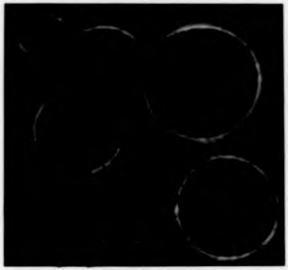


Figure P38

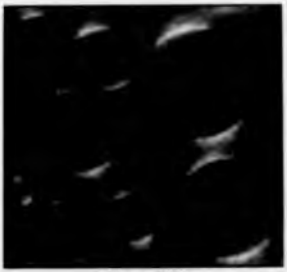


Figure P39

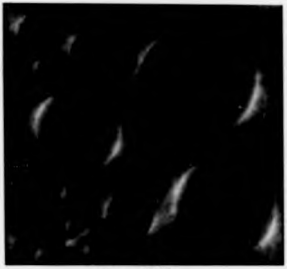


Figure P40

Figure P41

Figure P42

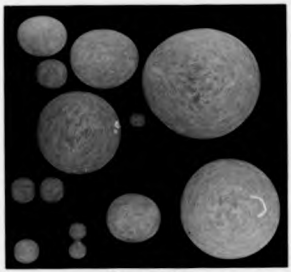


Figure P43

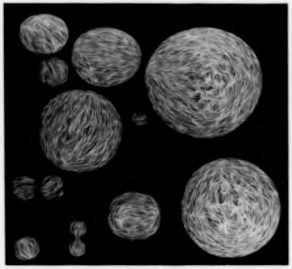


Figure P44

Figure P45

Figure P46

Figure P47

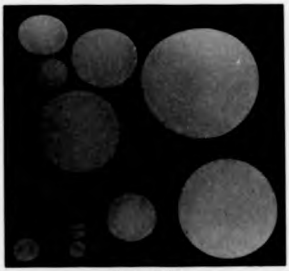


Figure P48

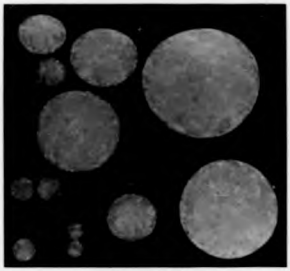


Figure P49

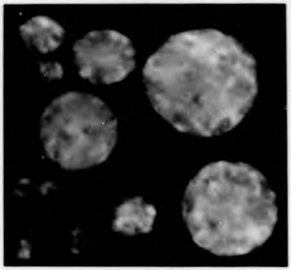


Figure P50

Figure P51

Figure P52

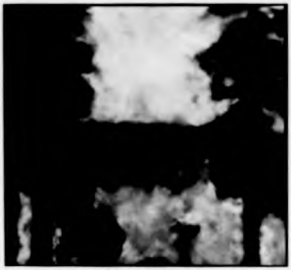


Figure P53

THE BRITISH LIBRARY DOCUMENT SUPPLY CENTRE

TITLE MULTIRESOLUTION IMAGE MODELLING AND ESTIMATION

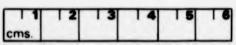
AUTHOR Simon Clippingdale B.Sc.

INSTITUTION and DATE 1988 The University of Warwick

Attention is drawn to the fact that the copyright of this thesis rests with its author.

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no information derived from it may be published without the author's prior written consent.

THE BRITISH LIBRARY
DOCUMENT SUPPLY CENTRE
Boston Spa, Wetherby
West Yorkshire
United Kingdom



REDUCTION X 20

CAMERA 8