Unsupervised Image Segmentation Combining Region and Boundary Estimation

Abhir Bhalerao
Roland Wilson
Department of Computer Science
University of Warwick
Coventry CV4 7AL
{abhir,rgw}@dcs.warwick.ac.uk

August 24, 2000

Abstract

An integrated approach to image segmentation is presented that combines region and boundary information using maximum *a posteriori* estimation and decision theory. The algorithm employs iterative, decision-directed estimation performed on a novel multiresolution representation. The use of a multiresolution technique ensures both robustness in noise and efficiency of computation, while the model-based estimation and decision process is flexible and spatially local, thus avoiding assumptions about global homogeneity or size and number of regions. A comparative evaluation of the method against region-only and boundary-only methods is presented and is shown to produce accurate segmentations at quite low signal-to-noise ratios.

Contents

1 Introduction **3**

2 A Two-stage, Region-Boundary Model for Segmentation **4**
2.1 Boundary Process 4
2.2 Region Process 5

3 The Segmentation Algorithm **6**

4 Experimental Results **15**

5 Conclusions **20**

1 Introduction

Image segmentation is one of the most widely studied problems in image analysis. The simplest techniques are based on a histogram of the data, which retains no spatial information. As a result, the classification is invariant to permutations of the pixels, ignoring the spatial coherence essential in segmentation [1]. The need to incorporate spatial information led to the development of methods using a *context* or neighbourhood [2, 3]. Markov Random Field (MRF) theory has been used to model the underlying spatial dependences [4, 5]. These models have been used in various maximum *a posteriori* (MAP) methods for image restoration and segmentation [6].

In problems such as segmentation, multiresolution analysis offers two key advantages over pixel-based methods:

1. It provides a way to trade-off class and spatial resolution; repeatedly blurring and subsampling the image decreases the noise and improves the class certainty, but at the expense of spatial resolution [1].

2. Solutions to optimisation problems at low resolutions tend to be close to the solution at high resolution because most of the variation in images is at scales which are large compared to 1 pixel. It is therefore possible to avoid local minima which lie far from the global minimum, without having to rely on stochastic algorithms.

For example, the split-and-merge techniques introduced by Chen and Pavlidis [7], and later developed in [8, 9, 10], use a linked pyramid and statistical decision criteria to combine global and local region information. The quadtree method of Spann and Wilson [11], and also in [12, 13, 14], used clustering of a historgram at a low spatial resolution, followed by boundary refinement. Recently, a multiresolution stochastic image model was used by Bouman and Shapiro in the development of a Bayesian segmentation algorithm[15]. Although region based methods have their attractions, it is quite hard to define region models which are sufficiently robust and general to allow 'just enough' intra-region variation, without causing ambiguities in what constitutes a region.

Correspondingly, the importance of line and edge information in both biological and computer vision systems has long been recognised. As in region growing, there is a fundamental trade-off in edge detection between spatial localisation and noise immunity, as shown by Canny [16]. The outputs from the edge detection must then be grouped into boundary curves, to complete the segmentation. Three ways of achieving this have been described: techniques such as the 'active contour' models, which rely on energy minimisation at pixel level [17], global methods, such as the Hough transform (HT) [18], or intermediate approaches, such as the hierarchical HT [19] and Multiresolution Fourier Transform (MFT) based methods [20].

It is clearly attractive to combine region processing with boundary detection to overcome their respective shortcomings. Haddon and Boyce extended relaxation techniques to combine edge and region information, using co-occurrence matrices [21]. In MRF modelling, edges are incorporated by the introduction of a binary valued line process [4, 22]. Although the results are generally better with a line process, the problem of parameter estimation for the joint model is compounded. Also, the size of the line process neighbourhood is tied to that of the gray level process, which is kept small to make the computation tractable. A more relevant example of combining a region based segmentation

with edge detection was presented by Pavlidis and Liow [23] and [24]. Recently, Zhu and Yuille reported an algorithm which is claimed to unify a number of current techniques [25]. Their approach, an iterative Bayesian method, alternates boundary estimation and region estimation steps. It requires the selection of a number of seed regions for initialisation of the statistical measurements on which the region estimation is based. It would be advantageous both to minimise dependence on such initial conditions and for the region and boundary processing to be autonomous, so that where necessary either could be used independently of the other.

This paper describes an algorithm combining region and boundary estimation using maximum *a posteriori* (MAP) estimation, which satisfies these requirements. It is based on co-operation between autonomous region and boundary processing; there is no dependence on seeds, nor is a merging step required. The algorithm employs iterative, decision-directed estimation performed on a multiresolution representation, which ensures both robustness in noise and efficiency. The model-based estimation process is adaptive and spatially local, avoiding assumptions about homogeneity or size and number of regions which characterise some earlier algorithms. A two stage combined region-boundary model for segmentation is introduced in section 2. Then, the algorithm is described in detail, followed by results on the segmentation of objects with complex shapes in a high level of noise (0 dB SNR) and on natural images. Comparative results are presented with a quad-tree region-only method [11], an MRF based technique Iterated Conditional Modes (ICM) [26] and on an implementation of the popular Canny edge detector [16].

2 A Two-stage, Region-Boundary Model for Segmentation

The image model is derived from one quad-tree image model introduced in [27], forms of which have been applied in the areas of image estimation [28], coding [29] and curve extraction [30]. The image can be visualised as being built up by successive approximation through scale, beginning with a coarse description at a low resolution and progressively refining it through scale. To introduce multiple regions, however, a more complex scheme is necessary: images are realised by a two-stage process. First, region *boundaries* (as connected piecewise linear curves) are defined by a randomly placed set of vertex points. Secondly, connected regions are modelled as an MRF [31, 4], where the adjacency of the pixels is constrained by the boundaries within the boundary process.

2.1 Boundary Process

The boundary model differs significantly from that employed in 'snakes' and similar 'active shape' or deformable models, in that it replaces smoothness constraints on the curve by a polygonal model defined by a set of vertex points, with spatial co-ordinates $\vec{\chi}_r = (p, q)$, and links between one or more points, λ_{rs}, forming a contour. The vertices may be modelled by any suitable point process in the plane [32], such as a Poisson process, with probability

$$P(n, A) = (\beta A)^n e^{-\beta A}/n! \tag{1}$$

that the contour contains exactly n vertices, $\vec{\chi}_r, 1 \leq r \leq n$, distributed randomly in an area A of the image. The parameter β defines the rate at which vertices appear in the image. This amounts to a constraint not on the *smoothness*, but the *complexity* of the boundary curve, which is fundamental in the segmentation algorithm: the best boundary is one with fewest boundary points, among those which approximate the data equally well. Note that by choosing more complex models than the Poisson process, more sophisticated and realistic boundaries are possible: (1) represents the minimal *a priori* constraint, yet provides a *complete* boundary representation, since any plane curve of finite length can be approximated arbitrarily well by a piecewise linear curve. These two aspects make the model particularly appropriate for MAP estimation, since the number of vertices is the key parameter defining both the prior probability, via (1) and the likelihood, i.e. the fit to the data.

2.2 Region Process

A quadtree structure is used in conjunction with the boundaries to define regions, with the spatial co-ordinates $\vec{\xi}_i(l) = (p, q, l)$ representing square regions $\Lambda_i(l)$ of size $2^{(l-1)}$. On any given level, l, these sets are clearly disjoint, and together define a conventional quadtree tessellation.

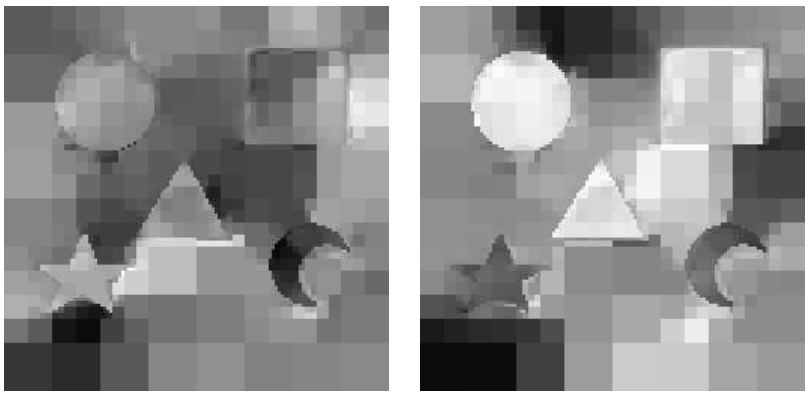
A coarse-to-fine refinement strategy is used to split each region block into 4 if the block contains one or more boundary links λ_{rs}. If there are no boundary features in a given block, then it is classed as being interior and the quadtree is truncated at that point (a leaf node). Correspondingly, at any level there are two classes of pixels: those that are interior pixels and those containing a boundary segment.

Within this irregular lattice structure, the image is modelled as a Gaussian MRF with zero mean [33, 34, 35, 36]. This requires a suitably defined set of neighbours $\mathcal{N}_i$, for each leaf of the quadtree $\vec{\xi}_i(l)$: two leaf vertices are linked if their corresponding image blocks $\Lambda_i(l)$ share an edge and no boundary link intersects the line joining their centres. This allows *region* links $\gamma_{ij} \in \{0, 1\}$ to be defined analogously to boundary links and hence a region graph. The neighbour set $\mathcal{N}_i$ is just the set of regions j for which $\gamma_{ij} = 1$.

Expressing the dependence of the value of node $\vec{\xi}_i(l)$ on its neighbours (the 2-dimensional Markov property [31, 26, 4])

$$\begin{aligned} P(x_i(l) \mid x_j(k), j \neq i) &= P(x_i(l) \mid x_j(k), j \in \mathcal{N}_i(l)) \\ &\propto \exp\{-\frac{1}{2\nu(l)}(x_i(l) - \sum_{\vec{\xi}_j(k) \in \mathcal{N}_i(l)} \beta_{ij} x_j(k))^2\} \end{aligned} \tag{2}$$

where $x_i(l)$ is the gray level associated with the leaf node $\vec{\xi}_i(l)$, β_{ij} are the MRF parameters and $\nu(l)$ are the conditional variances, which depend on the block size. In order to obtain a realisation at the image plane from the quadtree region model, an interpolation function is defined, which in the simplest case copies each node value $x_i(l)$ to each pixel in the corresponding block $\Lambda_i(l)$. Boundary pixels are simply interpolated from their neighbours. For some applications, the obvious block artefacts resulting from the quadtree interpolation may be a problem, but in the present application, it has not proved to be one.



(a) after 100 iterations (b) after 1000 iterations

Figure 1: Model realisations for regions derived from 'shapes' image. Note the different block tessellations between (a) and (b).

Examples of realisations of the model, based on links derived from the 'shapes' image are shown in figures 1(a) and 1(b). These show the results of 100 and 1000 iterations. Because the regions are separate, there is no leakage from one to another and so their gray levels evolve independently. This indicates that the model does indeed capture important properties of images.

3 The Segmentation Algorithm

Segmentation is based on the model of section 2, corrupted by additive white Gaussian noise, giving data which are jointly normal

$$y_{ij} = x_{ij} + \nu_{ij} \tag{3}$$

The algorithm consists of two main stages: an initialisation stage which selects the nodes of the region and boundary graphs $\{R_0, B_0\}$, followed by an iterative stage which aims to minimise the energy functions in each graph until convergence (Figure 2). The algorithm is based on the simplest multiresolution representation: the lowpass pyramid, [37] [38] [39] [40]. From the pyramid, a set of nodes is selected as forming an estimate of the model quadtree: $D = \{R_0, B_0\}$ (stage 1). Links to neighbours are then defined in

1. Initialise $D = \{R_0, B_0\}$
2. Repeat until converged
   (i) Update state of region graph R_t
   (ii) Update state of boundary graph B_t

Figure 2: The segmentation algorithm consists of two main stages: initialisation followed by iterative processing and cooperative interaction between R_t and B_t.

a modified form of a region adjacency graph [41] [42]. Each link has a binary state: on or off, which determines the membership of a node to a given processing neighbourhood. MAP tests are used in updating the link and node states (step 2(i)), and in node insertion/deletion. Each node has state variables that are updated in step 2: pixel grey level values and local variances for region nodes, and spatial position and edge strength, local orientation and corresponding variances for boundary nodes.

The interaction between the processes follows the model described in section 2 above, with the boundary state affecting the region state probabilities and vice versa. The overall

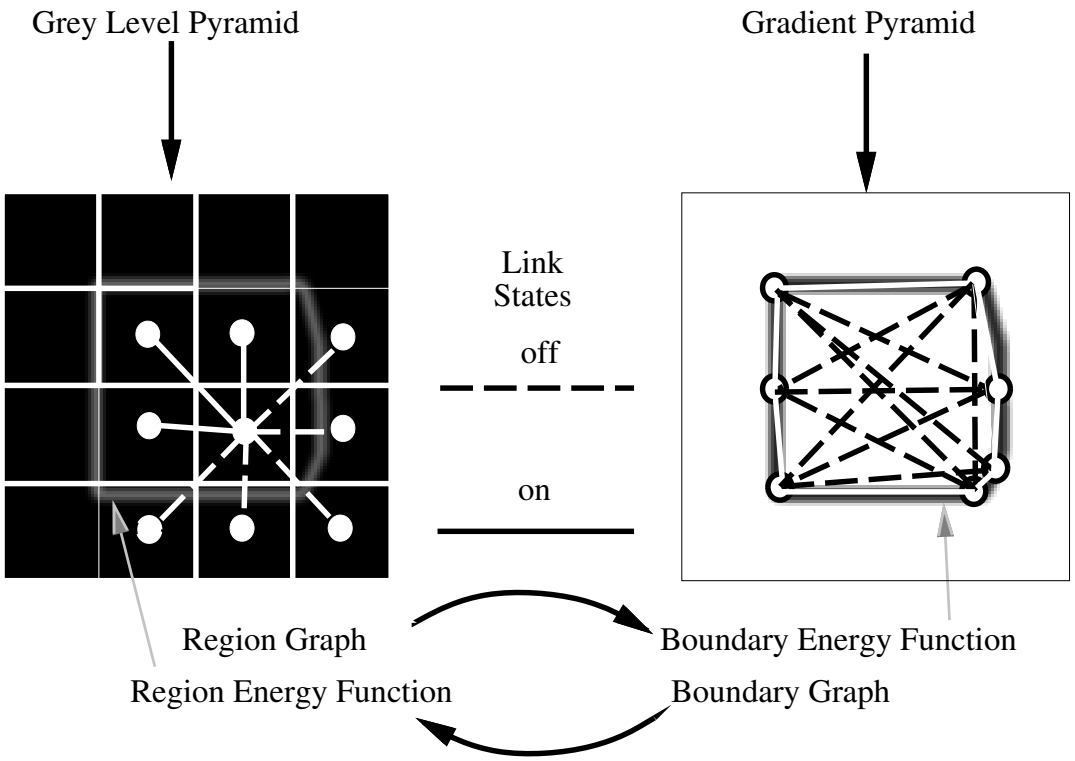


Figure 3: The region and boundary graphs interact via their energy (loglikelihood) functions. Each graph is moved towards a low energy (high probability) state iteratively.

structure is illustrated in figure 3, for a simple image containing a single shape. In effect, the algorithm changes the states of the two graphs to align both of them more closely with the energy landscape. The energy represents the edge loglikelihood and it therefore lies perpendicular to the region links and along the boundary links. Off links in the region graph, indicating the presence of an edge between adjacent region blocks, contribute a term to the boundary energy function. 'On' links in the boundary graph, where there is certainty of a boundary at the current interaction, influence the region link states.

Graph refinement takes places by node insertion and deletion after node values have been updated. Insertion of nodes into the region graph by block splitting parents into 4 children within the quadtree structure, focuses the edges. Conversely, node deletion in the boundary graph minimises the description length and redistributes boundary nodes to better fit the gradient energy landscape.

All parameter estimation is performed on-line: the algorithm is unsupervised and adaptive to the data. This is important if the processing is to be kept as general as possi-

ble. Statistics such as sample variances are gathered locally using the pyramid structure, allowing variation in statistical properties across the image: the estimates used in a given level of the pyramid come from the ancestor node four levels above the given node. The advantage of such a procedure is illustrated in figure 9, in which the gray level difference between regions varies considerably, but segmentation is still achieved.

In the following, the details of the algorithm are presented according to Figure 2. After a description of the initialisation stage, the iterative steps are described separately for the region and boundary processing. The section is concluded with details about the node/insertion deletion step, and a discussion on the convergence.

1. Initialisation

Although the segmentation algorithm is iterative, the initial state of the graphs can have a major impact on both the number of iterations to reach a steady state and, because the algorithm is deterministic, the quality of the final result. A robust initial estimate is therefore an important component of the algorithm.

The first step is a smoothing operation performed using a lowpass pyramid [40] which trades off noise reduction against spatial resolution. Given an image $y_{ij}, 0 \leq i,j < N = 2^M$, and a lowpass kernel $A_{mn}, -K \leq m,n < K$, the general form of the processing is

$$y_{ij}(l) = \sum_{m=-K}^{K} \sum_{n=-K}^{K} A_{mn} y_{(2i-m)(2j-n)}(l+1) \tag{4}$$

where $l_0 \leq l \leq M$ and $y_{ij}(M) = y_{ij}$. The kernel used in this work was designed by a least squares approach to minimise the effects of aliasing errors in gradient estimation, which forms a significant part of the overall algorithm [43]. The grey level pyramid is used directly in computing the initial region graph and forms the input to a gradient estimator, whose output is used to produce the initial boundary graph.

Region node selection: R_0

To initialise the region graph, leaf nodes must be selected from the quad-tree representation of the pyramid. This process can be made independent of the boundary computation, using the residual error obtained from prediction of level $l+1$ of the pyramid from level l (as in the Laplacian pyramid [38])

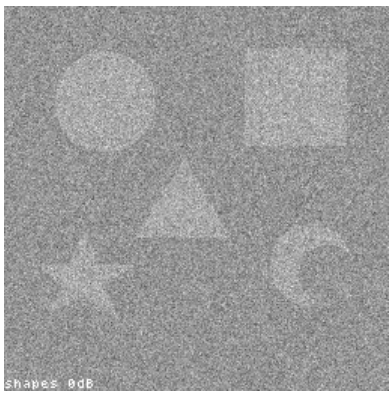
$$\hat{x}_{ij}(l+1) = \sum_{m,n} A_{mn} y_{(i/2-m)(j/2-n)}(l) \tag{5}$$

with a corresponding residual variance estimate

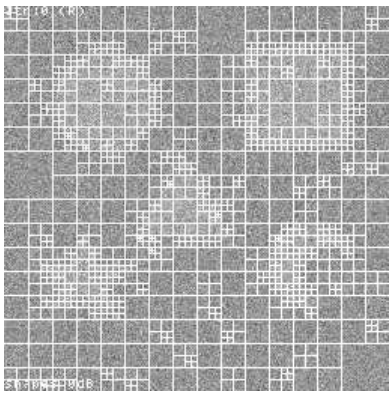
$$v_{ij}^2(l) = \sum_{m,n} A_{mn} (y_{(2i-m)(2j-n)}(l+1) - \hat{x}_{(2i-m)(2j-n)}(l+1))^2 \tag{6}$$

A simple and robust test statistic for the presence of edges within the block i,j,l is the normalised residual variance

$$\epsilon_{ij}(l) = \frac{v_{ij}^2(l)}{\bar{v}^2(l)} \tag{7}$$



(a) 'shapes' 0dB



(b) Node selection $R_0, \alpha(l) = 0.5$

Figure 4: Initialisation of the region ajacency graph R_0 prior to iterative processing

-0.074	-0.095	0.000
-0.095	0.000	0.095
0.000	0.095	0.074

g_0

0.000	0.095	0.074
-0.095	0.000	0.095
-0.074	-0.095	0.000

g_1

Table 1: 3×3 gradient edge kernels: g_0, g_1 used to derive orientation estimate $\theta(l)$ from grey-level pyramid.

where $\bar{v}^2(l)$ denotes sample variance, since nodes containing edges will have a significantly higher residual error than interior nodes. If $\epsilon_{ij}(l) > \alpha(l)$, the threshold on level l, the node i, j, l is marked; the lowest marked node (ie. highest l) in any subtree becomes a leaf of the quadtree region representation (cf. section 2), ie a vertex $\vec{\xi}_k$ of the region graph. Figure 4(b) shows the result of this process on the noisy 'shapes' image using $\alpha(l) = 0.5$. It can be seen that region interiors tend to contain large blocks, while smaller blocks dominate the boundaries. This allows the regions to 'grow' out from the more reliable 'core' of large blocks towards the edges, but avoids the well-documented problems which conventional multiresolution methods have with high curvature boundaries. Although the neighbourhood structure for the tree is more complex than a conventional image array, the processing overhead is more than offset by the relatively small number of leaf nodes (typically a few hundred).

Boundary node selection: B_0

The boundary process is controlled by a gradient field derived from the grey level pyramid. The initial gradient estimate is obtained by the use of a pair of optimised (3×3) spatial kernels shown in table 1 [40]. The gradient components, $g_k(l), 0 \leq k \leq 1$, are combined to yield an estimate of the local double-angle gradient vector $\theta_{ij}(l)$. The magnitude of this vector is a measure of the local edge energy and the argument *twice* the local gradient angle. The doubling of the the angle, originally proposed by Knutsson [44], is useful in local averaging and, more significantly, essential in representing the boundaries of vector data, such as colour or texture features, for which no full ordering exists.

Because of the noise-sensitivity of gradient estimates, smoothing is needed to produce usable data for the boundary detection. The principal dependences between gradient data

level	*Input SNR (dB)*	*Output SNR (dB)*
4	50.8	50.8
5	38.4	39.8
6	27.5	32.6
7	13.1	28.0
8	-4.4	21.9

Table 2: Vector estimation errors of gradient enhancement results on 'shapes' 0dB image. SNR of 32dB is approximately a 1 deg. error in angle.

are from two sources: the neighbours $\mathcal{N}_{ij}(l)$ of the ijth pixel on the same level of the pyramid; the 'ancestors' of the pixel on level $l-1$, $\mathcal{N}_{ij}(l-1)$,

$$P(\theta_{ij}(l) \mid \{\theta_{mn}(k), \forall m, n, k\}) = P(\theta_{ij}(l) \mid \{\theta_{mn}(k), k = l, (m,n) \in \mathcal{N}_{ij}(l); \\ k = l-1, (m,n) \in \mathcal{N}_{ij}(l-1)\}) \quad (8)$$

An efficient way to use these correlations is a combination of sequential estimation through scale [45], followed by iterative estimation on each scale, as outlined below:

For each level $l > l_0$,

(1) Propagate estimates $\theta_{ij}(l-1)$ from level $l-1$, using interpolator A_{mn}.

(2) Iterative estimation on level l, using an anisotropic filter $H_{mn}(\theta)$, whose orientation is controlled by the current estimate $\theta_{ij}(l)$.

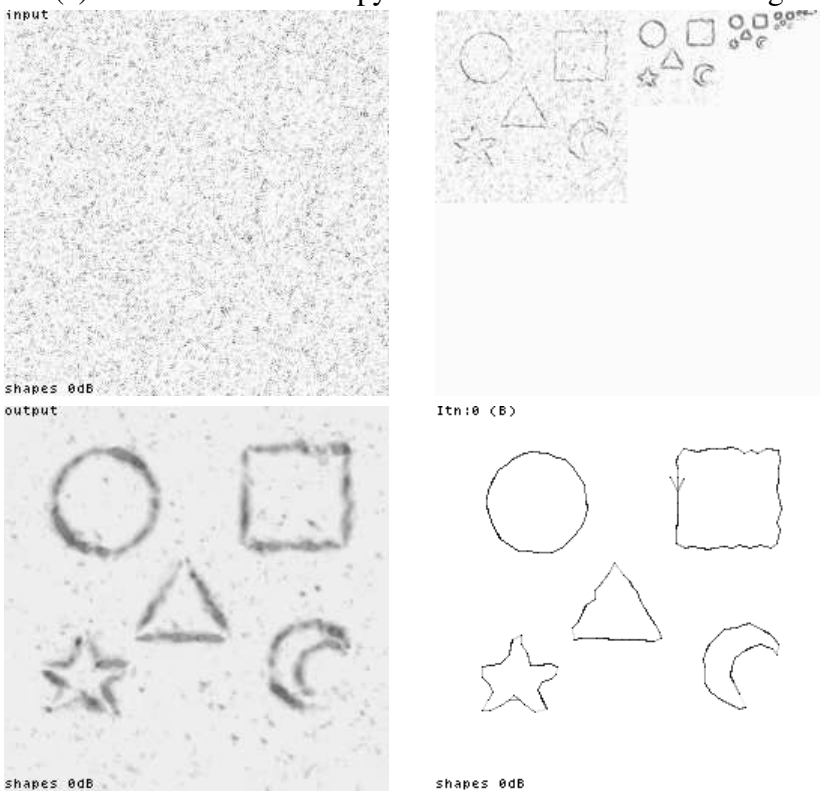
Details of the propagation coefficents, design of the anisotropic filter, $H(\theta)$ and associated parameter estimation are given in [46]. As an illustration, from the magnitude of the gradient pyramid of figure 5(a) there is clearly no reliable edge information on the lowest level of the pyramid. Inspection of the lower spatial resolutions, however, shows that the initial gradient estimation picks up the edges at these levels. The result of recursive filtering in the pyramid is apparent in the output pyramid (figure 5(b)). The vector estimation errors are given as a SNR of the recursive enhancement process are given in table 2, as a function of level, from the starting level (4) to full spatial resolution (level 8). Note that a 1 deg. error in angle corresponds to a SNR of 32dB. It can be seen that there is a significant improvement in SNR as a result of the processing.

A set of candidate boundary nodes $B = \{\vec{\chi}_0, \vec{\chi}_1, \ldots, \vec{\chi}_n\}$ is selected, using peak detection on the gradient pyramid. Selected nodes are refined by choosing among their children that node having the largest inner product of gradient vectors with the parent, ie. the quadtree node i, j, l maximising the likelihood

$$P(\theta_{ij}(l) \mid \theta_{i/2j/2}(l-1) = f(\theta_{ij}(l) - \theta_{i/2j/2}(l-1)) \quad (9)$$

where the density $f(.)$ is modelled as normal in the experiments. The initial boundary estimate, B_0, is shown in figure 5(c) for the example 'shapes' image.

(a) Gradient estimation pyramid before recursive filtering



(b) Final gradient estimate (c) Initial boundary estimate B_0.

Figure 5: (a) Initialisation of boundary graph B_0 after recursive smoothing of gradient pyramid of 'shapes' image 256×256. Note that at 0dB SNR, prior to recursive estimation, the gradient estimate at level 0 contains no valid edge information

2. Iterative Processing

2(i) Updating Region States

First, the states of the links between pairs of nodes are estimated using

$$\gamma_{ij}^t = \begin{cases} 1 & \text{if } P_{ij}(1)p(x_i^{t-1} - x_j^{t-1} \mid 1) > \\ & \quad P_{ij}(0)p(x_i^{t-1} - x_j^{t-1} \mid 0) \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

$P_{ij}(1)$ is the prior probability that the link between the nodes $\vec{\xi}_i$ and $\vec{\xi}_j$ is 'on' and depends on an estimate of the likelihood of an edge in the vicinity of node i, based on the gradient data θ.

$$P_{ij}(1) = P(\overline{edge}|\theta) = P(\overline{edge})p(\theta|\overline{edge}), \tag{11}$$

and similarly

$$P_{ij}(0) = P(edge|\theta) = P(edge)p(\theta|edge) \tag{12}$$

We can approximate the conditional densities $p(\theta|\overline{edge})$ and $p(\theta|edge)$ by assuming a normal model distribution for the gradient magnitude:

$$p(\theta|\overline{edge}) = \frac{1}{\sqrt{2\pi}\tau} \exp[-0.5(\theta_{max} - 0)^2/\tau^2)], \tag{13}$$

(the expected mean edge magnitude for block with no edges namely 0) and

$$p(\theta|edge) = \frac{1}{\sqrt{2\pi}\tau}, \tag{14}$$

this time taking θ_{max} to be the expected mean. The reason for using θ_{max} is that the mean magnitude, θ_{mean}, will be too small, close to the RMS gradient magnitude τ when estimated across all the pixels in a block. The prior $P(\text{edge})$ is the probability that a link crosses an edge (approximated from link intersections in the initial graph configurations). This provides a way for the boundary processing to influence the region linking. For the test in equation 10, it is sufficient to calculate the ratio $P_{ij}(1)/P_{ij}(0)$.

The conditional densities in equation 10 of the difference between the grey levels at two neighbouring nodes (i, j) under the two hypotheses:

(1) they belong to the same region and

(2) they don't belong to the same region and have independent means

Both tests, as with estimation of the link prior above, assume a normal model and are performed by considering their difference $x_i - x_j$ at iteration t against the variance of this statistic i.e. $E[(x_i - x_j)^2] = v_i^2 + v_j^2$. For test (1), if the nodes i and j belong to the same region, their mean difference will be zero and have a variance equal to the intra-region noise. For test (2) the nodes will have independent means whose difference can at approximated by the values at $t = 0$ i.e. $(x_i^0 - x_j^0)$. The normal likelihoods are explicitly calculated from:

$$\begin{aligned} p(x_i^t - x_j^t \mid 1) &\propto \exp\left(\frac{-(x_i^{t-1} - x_j^{t-1})^2}{2(\overline{v}_i^2 + \overline{v}_j^2)}\right) \\ p(x_i^t - x_j^t \mid 0) &\propto \exp\left(\frac{-((x_i^{t-1} - x_j^{t-1}) - (x_i^0 - x_j^0))^2}{2(\overline{v}_i^2 + \overline{v}_j^2)}\right) \end{aligned} \tag{15}$$

where x_i^t is the estimated grey level at the node i on iteration t and $\overline{v}_i^2$ is an estimate of the intra-region variance between nodes on the same level as $\vec{\xi}_i$.

Having obtained the link states, the node grey levels, x_i are updated. The estimation scheme shares some features with the Bayesian methods described in [4], but no annealing is used, and is closer to Besag's Iterated Conditional Modes (ICM) scheme [26] which employs a local probability maximisation. The main differences from ICM are the irregular graph and the dynamic control of neighbourhood sets by the link states, which are affected by the gradient data. Thus, from the link states γ_{ij}^t, the neighbourhood of node i is updated

$$\mathcal{N}_i^t = \{j : \gamma_{ij}^t = 1\} \tag{16}$$

The deterministic update rule amounts to an averaging in the neighbourhood $\mathcal{N}_i^t$:

$$x_i^t = w_i x_i^0 + \sum_{j \in \mathcal{N}_i^t} w_j x_j^{t-1} \tag{17}$$

where w_i are the model parameters, which for a piecewise constant region model are set proportional to the area represented by the node $w_i = |\Lambda_i(l)|$, corresponding to the inverse of the noise variance at that node. The first term in (17) represents the likelihood and the second the prior.

Since regions corresponding to quadtree nodes above image level may straddle a boundary, each such node with a significant probability of containing an edge is tested for splitting. Any node $\vec{\xi}_i$ having at least one 'off' link is replaced by its 4 quadtree children if

$$P_i(\text{split} \mid \{x_k, k \in \mathcal{C}_i\}) > P_i(\overline{\text{split}} \mid \{x_k, k \in \mathcal{C}_i\}) \tag{18}$$

$\mathcal{C}_i$ is the set of quadtree children of $\vec{\xi}_i$. The prior $P_i(\text{split})$ is derived from the gradient data in a similar fashion to that for the boundary links: it is the posterior probability that at least one boundary line $\mathcal{L}_{kl}$ intersects the block $\vec{\xi}_i$

$$P_i(\text{split}) = \max_{(k,l) \in \mathcal{B}_i} P(\lambda_{kl} \mid \{\theta_{mn}(M)\}) \tag{19}$$

where $\mathcal{B}_i$ is the set of boundary vertices $\vec{\chi}_j$ whose links intersect the quadtree region defined by $\vec{\xi}_i$ and the conditional is the posterior probability that the boundary link between $\vec{\chi}_k$ and $\vec{\chi}_l$ is 'on'.

2(ii) Updating Boundary State

There are two sources of boundary information: the gradient estimates from the filtered gradient data, $\theta^B(M)$, and the region link and node states at iteration t. A simple way to include the latter information in the boundary state is to infer a gradient estimate from the region state. This is done by making a piecewise linear approximation of an edge for any region link which is 'off' at time t, ie. $\gamma_{ij}^t = 0$, using the gray level estimates x_i^t. This gives a 'prior' for the gradient θ^R, which can be optimally combined with the data to give an estimate at iteration t by the scalar Kalman filter update equation [47]:

$$\theta^t = \theta^{Rt} + \alpha^t(\theta^B - \theta^{Rt}) \tag{20}$$

where $(\theta^B - \theta^{Rt})$ is the prediction error given the observation and the gain α^t is calculated from the (estimated) variances of the two sources:

$$\alpha^t = \frac{var[\theta^{Rt}]}{var[\theta^{Rt}] + var[\theta^B]} \tag{21}$$

As the number of iterations increases, the region quadtree is refined and so the estimates based on the region states become more reliable. The result is that more weight is given to them, moving the boundary towards duality with the region graph.

Boundary links are set using a similar criterion to the region links, from the smoothed gradient data at full image resolution, θ^t_{mn} from (20). If $\mathcal{L}_{ij} = \{\alpha\vec{\chi}_i + (1-\alpha)\vec{\chi}_j, 0 \leq \alpha \leq 1\}$ is the line joining boundary vertices $\vec{\chi}_i$ and $\vec{\chi}_j$, the vector

$$\theta_{ij} = \frac{1}{|\mathcal{L}_{ij}|} \sum_{(m,n)\in\mathcal{L}_{ij}} \theta_{mn}(M) \tag{22}$$

represents the normalised gradient data along the line. Then the posterior probability for the link $\lambda_{ij} = 1$ is

$$P(\lambda_{ij} = 1 \mid \theta_{ij}) \propto P_\lambda(\|\vec{\chi}_i - \vec{\chi}_j\|) f(\theta_{ij} - \phi_{ij}) \tag{23}$$

where $P_\lambda(r)$ is the prior probability that two nodes r pixels apart are linked and ϕ_{ij} is a vector of unit length in the direction of the line $\mathcal{L}_{ij}$. For the experiments reported below, a normal model of the gradient data is appropriate and the prior was uniform. The variance of the gradient data was estimated locally using quadtree smoothing, as for the other data ([46]).

Vertices are constrained to be of degree ≤ 2 by selecting the two nodes j, k with maximum link probabilities as neighbours, if more than two pass the linking test of (23). As for region nodes, the neighbours of a boundary vertex are those to which it is linked at time t

$$\mathcal{N}_i = \{j, \lambda_{ij} = 1\} \tag{24}$$

The result of the link processing is a set of connected sets of vertices, each of which represents the boundary of a distinct object (see eg. figure 7(c)). As in most *active contour* models, links are not changed during the update cycle, as the adaptation of the model to the data is accomplished by operations on the vertices, as explained below.

To adapt the model to the gradient data, each boundary vertex $\vec{\chi}_i$ is moved to maximise the probability

$$P(\vec{\chi}_i \mid \{\theta_{mn}(M), (m,n) \in \mathcal{L}_{ij}, j \in \mathcal{N}_i\}) \propto P(\vec{\chi}_i) \prod_{j\in\mathcal{N}_i} f(\theta_{ij} - \phi_{ij}) \tag{25}$$

The prior $P(.)$ is uniform and the density $f(.)$ is $2-D$ normal, as in (23) above. The only way to implement this maximisation is via a local search.

At the same time, to take account of the prior, deletions and insertions into the list of boundary vertices are made by comparing the probabilities

$$P(\lambda_{jk} \mid \{\theta_{mn}(M)\}) \stackrel{?}{>} P(N+1 \mid N) P(\lambda_{ij} \mid \{\theta_{mn}(M)\}) P(\lambda_{ik} \mid \{\theta_{mn}(M)\}) \tag{26}$$

where N is the number of vertices excluding $\vec{\chi}_i$ and $P(N+1 \mid N) = e^{-\beta A}/(N+1)$ gives a Poisson distribution for the number of vertices. Approximate collinearity of the two links to neighbours is used to select vertices for deletion and new candidates for insertion are selected from local maxima in the gradient magnitude.

Convergence

Convergence is deemed to have occurred when there is no significant change in the grey level estimate between iterations *and* no significant movement of boundary nodes (less than 1 in the mean grey level, boundary vertex positional error of less than 0.5 pixel). This can typically be achieved in 10-20 iterations, the small number reflecting both the relatively small number of nodes in both boundary and region representations and the reliability of the initial estimates. The convergence characteristics thus compare very favourably with stochastic relaxation methods, e.g. simulated annealing, which often require hundreds of iterations, and the behaviour is more consistent with that of ICM. Figure 6 demonstrates the rapid convergence by showing the percentages region-link state changes per iterations for processing on the figure 4(a).

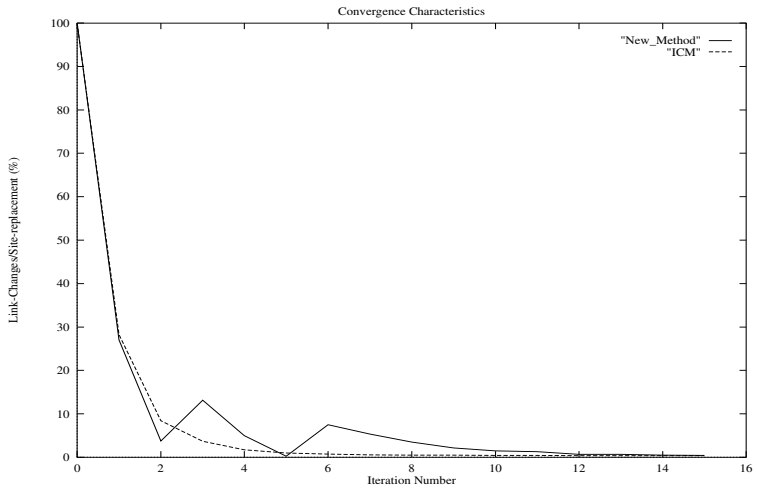


Figure 6: Percentages link changes per iteration for 'shapes' 0dB image compared with site-replacements using ICM illustrating similar convergence characteristics.

4 Experimental Results

Segmentation experiments were performed on both synthetic (simple object/background) and natural images. The images were chosen to highlight the following aspects of the new method:

- performance on simple shapes and those with sharp features (corners) (figure 7),
- performance on complex boundary shape (figure 9(a)-(c)),
- performance at low inter-class contrast in the presence of noise (graphs 10, 11),

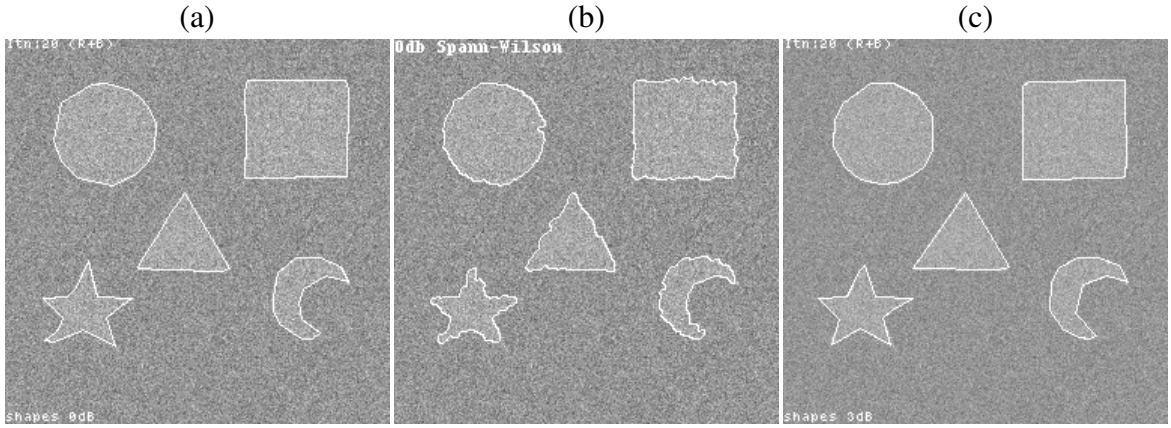


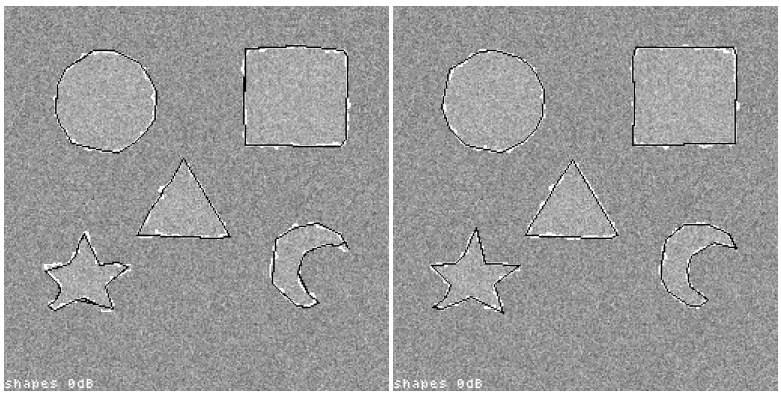
Figure 7: Results on 'shapes' image 256×256: (a) Boundary result on 0dB image, (b) Spann-Wilson result on 0dB image, (c) Boundary result on 3dB image.

- ability to locally adapt to variable signal-to-noise (figure 9(a)-(c)),
- ability to separate objects of varying size and close proximity to each other (figure 9(e)-(f)), and
- performance on natural images with within region grey level variation (figure 12).

Additionally, comparative results are presented using implementations of a quad-tree technique ([11]), and supervised region-only (MAP estimation using ICM [26], figure 14) and boundary-only methods (Canny edge detection [48], figure 13).

The first set of tests was carried out on the synthetic 256×256 'shapes' image shown in figure 4(a) (the initial processing from which the region and boundary graphs are obtained was illustrated in figure 4(b) and figure 5). The luminance difference between the objects and background is 20, as is the standard deviation of the added white Gaussian noise, giving an inter-region signal to noise ratio $20/\sigma = 1.0$. Figure 7(a) shows the final segmentation result on this test image. The segmentation of the circle compares well with the result obtained by the algorithm of Spann and Wilson [11](figure 7(b)), while performance on the objects that have sharp corner features is superior to those reported elsewhere. Note also that the vertices in the boundary graph do tend towards the vertices in the polygonal shapes. Figure 7(c) show an almost perfect segmentation result on a 3dB ($\sigma = 14$) version of the 'shapes'. These results show the boundaries superposed on the noisy input images. The boundary results illustrate the usefulness of a boundary model which can accommodate global structure: no pixel model could give such good estimates from the data. Figures 7(a) and (b) compare the boundaries defined by 'off' links between regions with those from the boundary graph, (a) without and (b) with interaction between the region and boundary processes. They show that the interaction is effective, both in moving the two graphs towards a state of duality and in reducing errors: without interaction, the region error is some 23% worse and the boundary error 14% worse than with interaction.

To test the method on more complex shapes, it was applied to the 'widgets' image (figure 9(a)). This image contains 4 objects with irregular boundaries that have been generated by a random-walk process. To this noise has been added to give SNR's of 12dB,



(a) (b)

Figure 8: Results of iterative processing without and with region-boundary interation: (a) Region off links (white) boundaries (black) without interaction, (b) Region off links and boundaries with interaction.

9.5dB, 6dB and 0dB in the same image, to provide a test of the local adaptivity of the method. The result of the region segmentation is shown in figure 9(b); white pixels indicate inter-region links which are 'off'. The corresponding boundary image (figure 9(c)) shows that, at the lowest SNR, the boundary is no longer closed, but this has not prevented the region process from separating the figure from the background. Figure 9(f) shows the results on the 'blobs' image at 0dB. This image has regions which are of varying sizes and some have been deliberately located close together to test the separation abilities of the segmentation. The resulting boundaries show that the segmentation method can cope well with the varying size, but not so well with the regions being close together. The boundary has been linked between some circles because of the unreliability of the local gradient information. The region result of figure 9(e) does not suffer from these defects. These results indicate that while boundary information is helpful in many cases, it is not sufficient in all cases: in both of these examples, the region process, guided by the boundary process, has been successful in identifying closed regions when the boundary estimate is erroneous.

The quantitative performance of the method over a variety of input SNR's is shown in figures 10 and 11. The boundary-error graph shows the root mean square (RMS) boundary displacement. Gradient estimation from the original 'shapes' test image was used in the comparison. The dashed lines show the comparative results for an implementation of the Spann and Wilson method [11] and a Canny edge-detector [48] with a hysteresis based thresholding. The parameters of the Canny implementation were chosen separately for each input image to produce the most pleasing qualitative results. Examples of the Canny implementation used are presented in figure 13. These figures illustrate well the compromise between having just enough smoothing to reduce the number of false edges and completing the object boundaries, whilst preserving the complexities in the boundary shape.

Figure 11 shows a comparison of the pixel-classification error calculated by counting the number of correctly classified pixels as a percentage of the total number of image pixels. Comparison is also made against the quad-tree method from [11], and a 'region-only' MAP estimation using ICM with a simple pair-wise interaction prior [26]. An ML

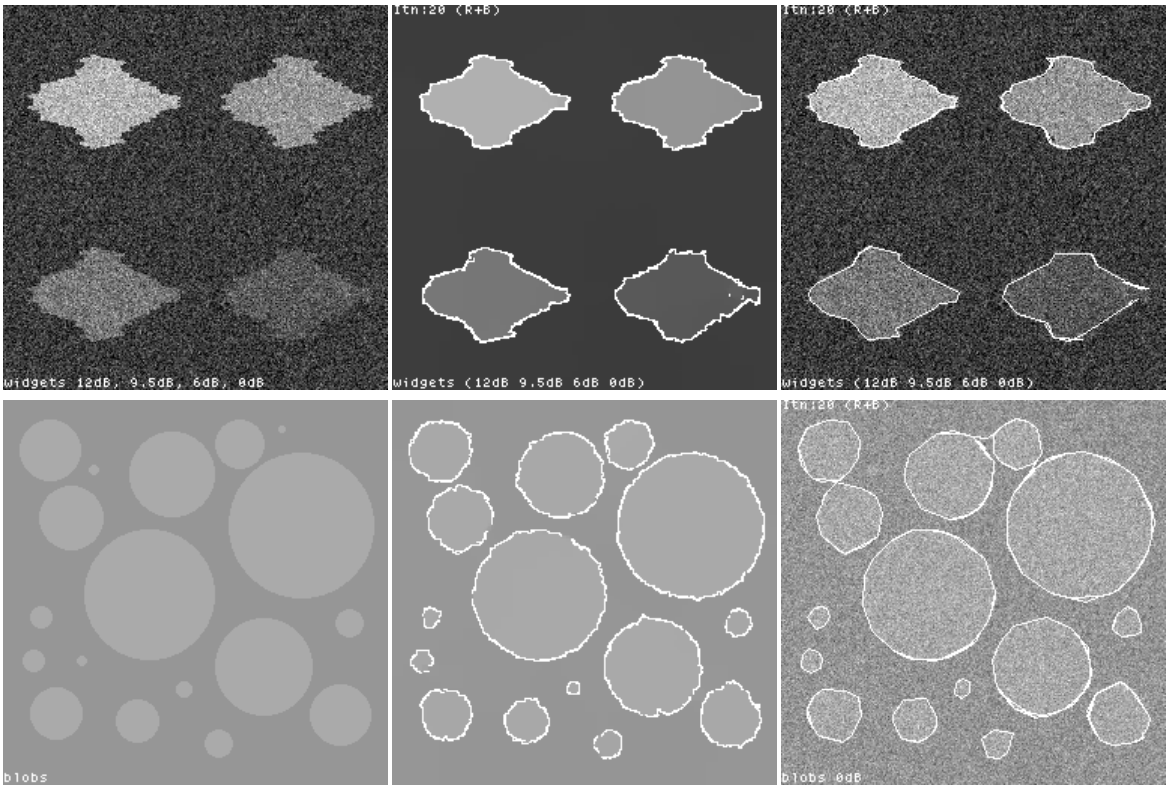


Figure 9: Results on ‘widgets’ and ‘blobs’ image: (a) ’widgets’ image (12dB,9.5dB,6dB,0dB), (b) Region result on ‘widgets’ image, (c) Boundary result overlayed on original, (d) ’blobs’ image (noise-free), (e) Region result on ‘blobs’ image, (f) Boundary result. Region results have ‘off’ links highlighted.

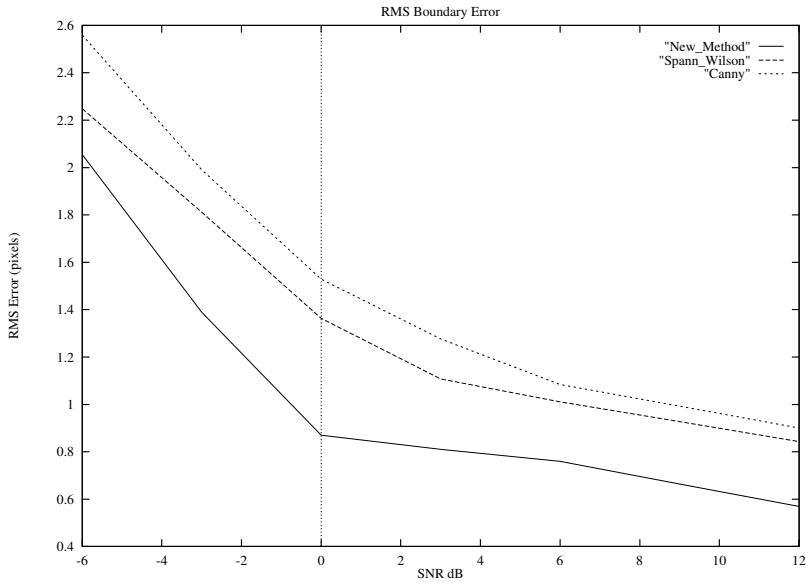


Figure 10: Comparative results on a set of progressively degraded images showing RMS boundary error in pixels (— New method, – – Spann-Wilson, - - - Canny). Boundary error is estimated as the perpendicular distance from the true object boundary in the noise free image. Errors below -3dB from Canny estimates are unreliable because of numerous false-edges in results.

estimate was used as the starting point for the ICM with all class means and noise variance given to the method a priori. The performance of two unsupervised multi-resolution methods are boardly similar, but the accuracy of ICM, which is critically reliant on a good initial estimate, rapidly deteriorates below 0dB. The reasons for poorer performance of our algorithm at SNR$>$ 0 is because our boundary model is piece wise linear and can only approximate curves in the test image (the circle and and crescent). Note however that the ICM segmentation is supervised and one would expect it to perform well in relatively noise free images. Indeed, it is likely that a less sophisticated segmentation approach than ours would be used on noise-free images. Some example result images of ICM are presented in 14 showing the initial ML estimate, and estimates at iterations 5 and 15. The small wrongly classified regions within the objects in figure 14(f) suggest that the optimisation is indeed trapped in a local minimum.

The results on two natural images are shown in figure 12(a)-(c) and figure 12(d)-(f). The ‘table’ image was chosen because it has many regions which are essentially flat and boundary shapes which are mostly polygonal, which conforms reasonably well to the image model. The region result of figure 12(b) shows that the principal surfaces have been identified and although there are a few artefacts caused by shadows in the region segmentation, these are not present in the boundary segmentation, but the main boundary curves have been detected (figure 12(c)). The ‘Lena’ image of figure 12(d) is more complex than the ‘table’ image, but still gives good boundary estimates, as can be seen from the boundary result in figure 12(f): although there are problems in textured and high curvature regions, such as the eyes, all of the significant boundaries have been extracted and linked. Because the region model is not adequate for the curved surfaces in the image, its success is limited, as can be seen from figure 12(e). Note, for example,

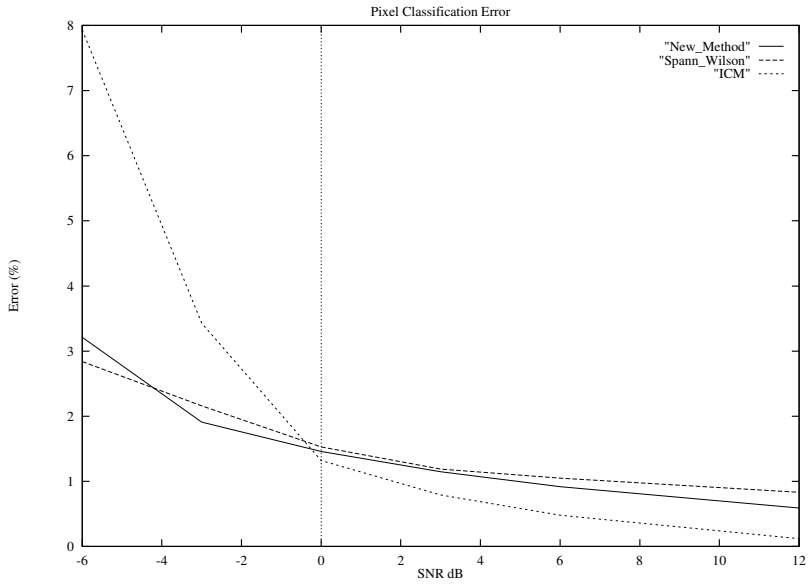


Figure 11: Comparative results on a set of progressively degraded images showing % pixel classification error. (— New method, – – Spann-Wilson, - - - ICM). Error is calculated as number classified pixels over total number of pixels. Note that ICM is supervised (means of objects and noise are given a priori). A maximum-likelihood estimate was used as initialisation for ICM [26].

that although the boundary around the top of the hat is closed, there are no 'off' links between the two regions for part of the boundary. This is easily explained by reference to the original image: there is no significant luminance edge at the very top of the hat and so the boundary can only be inferred, as it has been by the boundary process. The evidence is not strong enough to 'persuade' the region process, however. There are places in this image where neither process detects boundaries, for example between the hat rim and the forehead, again because of the lack of a luminance discontinuity, compounded by a complex arrangement of features. As noted in the introduction, these problems are a traditional weakness of region-based segmentation: they are a principal reason for using an explicit boundary representation in the present work. Nonetheless, the region process has contributed to the boundary identification, giving a result which compares favourably with Canny edge detection (e.g. figures 13(e) and (f)) and more sophisticated methods reported elsewhere (e.g. [20]).

5 Conclusions

A new model and algorithm for image segmentation have been described and been shown to be effective in describing and estimating regions and their boundaries with high accuracy from noisy data. In applications where region shape is important, it has a number of advantages over previously reported methods. It is computationally fast, flexible with respect to region and boundary models and uses only local processing. There are a number of features of the segmentation method presented here which differentiate it from work reported elsewhere (eg. [23], [15], [25]). The inhomogeneous block tessellation used by

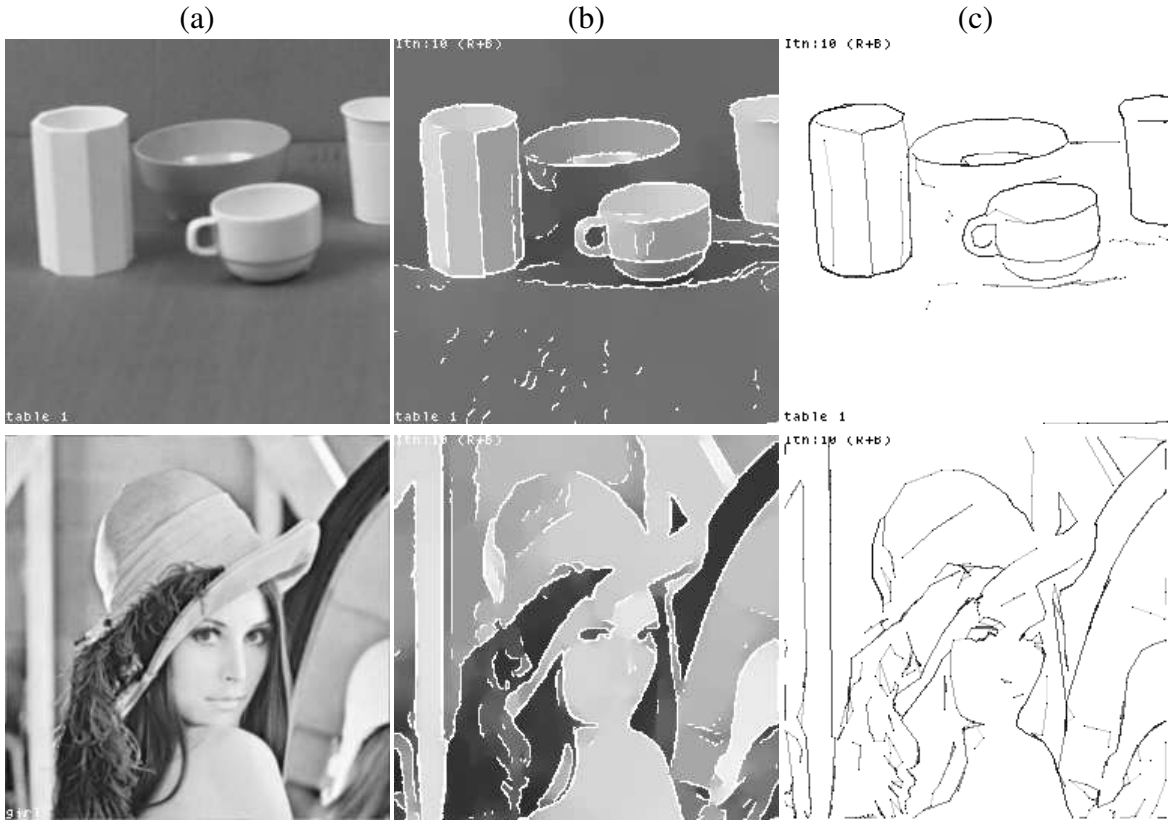
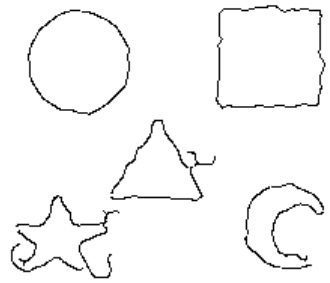
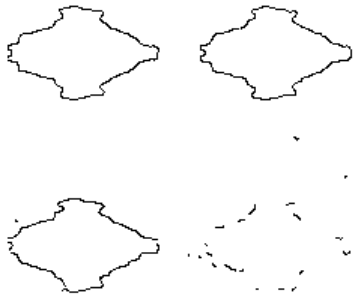


Figure 12: Results on natural images ‘table’ and ‘Lena’ image. (a) ‘table 1’ original, (b) Regions result, (c) Boundary result, (d) ‘Lena’ image, (e) Regions result, (c) Boundary result. ‘Off’ links highlighted in (b) and (e).

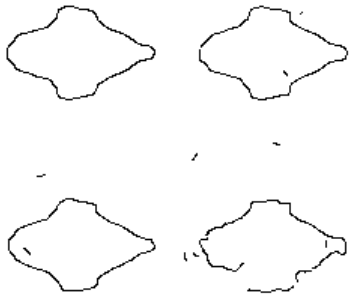
(a) $s = 1.75, low = 0.9, high = 0.92$



(b) $s = 3.00, low = 0.2, high = 0.9$



(c) $s = 1.50, low = 0.9, high = 0.95$



(d) $s = 2.70, low = 0.85, high = 0.9$

(e) $s = 1.50, low = 0.8, high = 0.85$

(f) $s = 1.50, low = 0.2, high = 0.6$

Figure 13: Comparative results using Canny edge detector with non-maximal suppression and hyteresis thresholding. Parameters used: $s =$ SD of separable Gaussian and hysteresis thersholds $\{low, high\}$ [48]. (All parameter values selected by visual inspection).

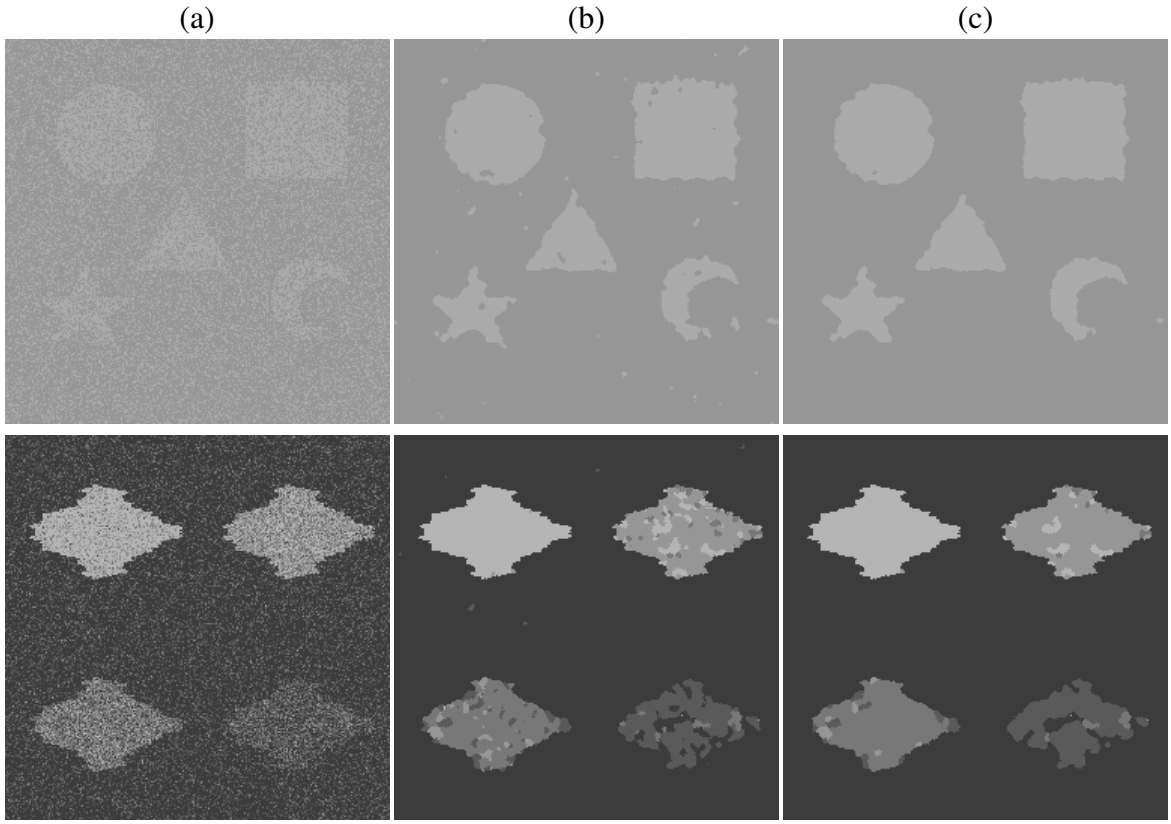


Figure 14: Results of running (supervised) Iterated Conditional Modes (ICM) [26] with a simple pair-wise interaction MRF on 'shapes' 0dB and noisy 'widgets' images. (a) and (d) are the initialisation step of ML classification, (b) and (e) after 5 iterations of ICM, (c) and (f) after 15 iterations. The increasing inter-class noise in (f) has resulted in many local minima in the final estimate. Compare (f) with figure 9(b).

the region estimation process, considerably reduces the computational demands of the iterative estimation. More significantly, it uses the most reliable data for its starting point, so that certainty about region membership grows out from a 'core' of larger interior blocks to the smaller boundary blocks as the estimation proceeds. Because of this, it is possible to avoid the substantial overhead of a stochastic algorithm, without sacrificing the quality of the result. Similarly, the polygonal boundary representation is general, compact and reliably estimated from the data. The graph representations of both regions and boundaries are, moreover, convenient for subsequent processing, to measure areas or perform shape recognition, for example. All parameter estimation is performed within the pyramid structure, making the method unsupervised and able to deal with arbitrary numbers of regions. The gradient estimation process is computationally cheap and robust. The efficiency of this estimator is partly a consequence of the use of a set of small filtering kernels designed specifically for pyramid construction and gradient estimation [43]. Finally, local orientation estimation is used as the vehicle for co-operation, which drives the region and boundary graphs towards duality. The experimental results have demonstrated the utility of this approach, with examples illustrating how the two processes together do a better job than either can alone.

There are several areas, however, in which consolidation and extension are required. Estimation of the model coefficients for each region, which are required for the iterative estimation (17), is a difficult problem, which has not been adequately addressed yet: the piecewise constant model used in the experiments can cope with some variation in grey level across regions, but is limited. One way to extend the method from grey level to other region properties, such as colour or local variance, is to make the region estimator multivariate. The algorithm in its present form is unsuitable for texture segmentation. However, as Spann and Wilson [1] show, it is not difficult to extend it to any region measurement, including colour or texture, simply by replacing the grey level by the appropriate (vector) of measurements, e.g. local variance, spectral properties or some other discriminator [49]. This could be easily accommodated within the overall structure, including the boundary estimation because it uses double angles ([44]).

Extension of the model to 3D imagery (from quad-tree to oct-tree) is also feasible, but in this case, the boundary model becomes more complicated having to take into account linear and planar gradient features [50]. The model for boundaries could also be extended to allow segments between nodes to be curved, i.e. quadratic arc segments or splines, which would result in a better fit to image curves. Moreover, in some applications, the two-connectedness of the boundary vertices is an unreasonable model, which could be replaced by one penalising, but not preventing, vertices of degree> 2. Finally, it may be of interest in some applications that the region and boundary representation allows the segmentation to be taken to sub-pixel resolution. These areas are the subject of continuing investigation.

References

[1] R. G. Wilson and M. Spann. *Image Segmentation and Uncertainty*. Pattern Recognition and Image Processing Series. Research Studies Press Ltd, 1988.

[2] R. A. Hummel and S. W. Zucker. On the Foundations of Relaxation Labeling Processes. *IEEE Trans. PAMI.*, 5(36):267–287, 1983.

[3] J. Kittler and J. Föglein. A General Contextual Classification Method for Segmentation. In *Proc. 3th Scandinavian Conf. Im. Analysis*, pages 90–95, Copenhagen, Denmark, 1983.

[4] S. Geman and D. Geman. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Trans. PAMI.*, 6(6):721–741, 1984.

[5] J. Besag. Towards Bayesian Image Analysis. In J. G. McWhirter, editor, *Mathematics In Signal Processing II*. Clarendon Press, Oxford, 1990.

[6] R. Chellappa and A. Jain, editors. *Markov Random Fields: Theory and Applications*. Academic Press, 1993.

[7] P. C. Chen and T. Pavlidis. Image Segmentation as an Estimation Problem. In A. Rosenfeld, editor, *Image Modelling*, pages 9–28. Academic Press, 1981.

[8] P. J. Burt, T. H. Hong, and A. Rosenfeld. Segmentation and Estimation of Image Region Properties Through Cooperative Hierarchical Computation. *IEEE Trans. Sys, Man Cyber.*, 11(12):802–809, 1981.

[9] T. H. Hong and A. Rosenfeld. Compact Region Extraction Using Weighted Pixel Linking in a Pyramid. *IEEE Trans. PAMI.*, 2(12):222–29, 1984.

[10] T. H. Hong and M. Shneier. Extracting Compact Objects Using Linked Pyramids. *IEEE Trans. PAMI.*, 6:229–237, 1984.

[11] M. Spann and R. G. Wilson. A Quad-Tree Approach to Image Segmentation Which Combines Statistical and Spatial Information. *Pattern Recognition*, 18(3/4):257–269, 1985.

[12] R. Wilson and M. Spann. Finite Prolate Spheroidial Sequence and Their Applications I, II: Image Feature Description and Segmentation. *IEEE Trans. PAMI.*, 10:193–203, 1988.

[13] J. Bigün. Frequency and Orientation Selective Texture Measures Using Linear Symmetry and Laplacian Pyramid. In *Proc. SPIE Conf. Visual Comms. and Image Processing*, pages 1319–1331, Lausanne, Switzerland, 1990.

[14] M. Spann and C. Horne. Image Segmentation Using a Dynamic Thresholding Pyramid. *Pattern Recognition*, 22(6):729–732, 1989.

[15] C. Bouman and B. Liu. Multiple Resolution Segmentation of Textured Images. *IEEE Trans. PAMI.*, 13:99–113, 1991.

[16] J. Canny. A Computational Approach to Edge Detection. *IEEE Trans. PAMI.*, 8:679–698, 1986.

[17] A. Witkin M. Kass and D. Terzopoulos. Snakes:Active Contour Models. *Intl. J. Computer Vision*, 1:321–331, 1987.

[18] R. O. Duda and P. E. Hart. Use of the Hough Transform to Detect Lines and Curves in Pictures. *Communications of the ACM*, 15:11–15, 1972.

[19] J. Princen, J. Illingworth, and J. Kittler. A Hierarchical Approach to Line Extraction Based on the Hough Transform. *Computer Vision, Graphics and Image Processing*, 52:57–77, 1990.

[20] R. Wilson, A. D. Calway, and E. R. S. Pearson. A Generalized Wavelet Transform for Fourier Analysis: the Multiresolution Fourier Transform and its Application to Image and Audio Signal Analysis. *IEEE Trans. IT, Special Issue on Wavelet Representations*, 1992.

[21] J. F. Haddon and J. F. Boyce. Simultaneous Image Segmentation and Edge Detection. In *Proc. IEE 3rd Int. Conf. on Image Proc. and its Applications*, pages 411–415, University of Warwick, U. K., 1989.

[22] D. Geman, S. Geman, C. Graffigne, and P. Dong. Boundary Detection by Constrained Optimisation. *IEEE Trans. PAMI.*, 12:609–28, 1990.

[23] T. Pavlidis and Y. T. Liow. Integrating Region Growing and Edge Detection. Technical Report TR.89.01.05, Department of Computer Science, State University of New York, Stony Brook, 1989.

[24] P. Bonnin and B. Zavidovique. An Edge PointRegion Cooperative Segmentation Specific to 3D Scene Reconstruction Application. In *Proc. SPIE Conf. Visual Comms. and Image Processing*, pages 1270–1281, Lausanne, Switzerland, 1990.

[25] S. C. Zhu and A. Yuille. Region Competition: Unifying Snakes, Region Growing and Bayes/MDL for Multiband Image Segmentation. *IEEE Trans. PAMI.*, 18:884–900, 1996.

[26] J. Besag. On the Statistical Analysis of Dirty Pictures. *Journal of the Royal Statistical Society*, 48(3):259–302, 1986.

[27] S. C. Clippingdale and R. Wilson. Quad-Tree Image Estimation: A New Image Model and its Application to MMSE Image Restoration. In *Proc. 5th Scandinavian Conf. Image Analysis*, pages 699–706, Stockholm, Sweden, 1987.

[28] S. C. Clippingdale and R. Wilson. Least Squares Estimation on a Multiresolution Pyramid. In *Proc. ICASSP-89*, pages 1409–12, Glasgow, 1989.

[29] M. Todd. *Image Data Compression Based on a Multiresolution Signal Model*. PhD thesis, Department of Computer Science, The University of Warwick, UK, November 1989.

[30] A. Calway. *The Multiresolution Fourier Transform: A general Purpose Tool for Image Analysis*. PhD thesis, Department of Computer Science, The University of Warwick, UK, September 1989.

[31] J. Besag. Spatial Interaction and the Statistical Analysis of Lattice Systems (with discussion). *Journal of the Royal Statistical Society*, 36(Series B):192–236, 1974.

[32] A.J. Baddesley and M.N.M Van Lieshout. Stochastic Geometry Models in High-level Vision. *J. Applied Statistics*, 20:231–56, 1993.

[33] J. W. Woods. Two-Dimensional Discrete Markovian Fields. *IEEE Trans. Information Theory*, 18:232–240, 1972.

[34] R. Chellappa and R. L. Kashyap. Digital Image Restoration using Spatial Interaction Models. *IEEE Trans. Acous. Speech Sig. Proc.*, 30:461–472, 1982.

[35] T. Simchony, R. Chellappa, and Z. Lichtenstein. Graduated Nonconvexity Algorithm for Image Estimation Using Computed Gauss Markov Field Models. In *Proc. ICASSP-89*, pages 1417–1420, Glasgow, 1989.

[36] P. W. Fung, G. Grebbin, and Y. Attikiouzel. Contextual Classification and Segmentation of Textured Images. In *Proc. ICASSP-90*, pages 2329–2332, Albuquerque, 1990.

[37] S. L. Tanimoto and T. Pavlidis. A Hierarchical Data Structure for Picture Processing. *Computer Vision, Graphics and Image Processing*, pages 104–119, 1975.

[38] P. J. Burt and E. H. Adelson. The Laplacian Pyramid as a Compact Image Code. *IEEE Trans. Comp.*, COM-31:532–540, 1983.

[39] P. Meer, E. S. Baugher, and A. Rosenfeld. Frequency Domain Analysis and Synthesis of Pyramid Generating Kernels. *IEEE Trans. PAMI.*, 9:512–22, 1987.

[40] R. Wilson and A. H. Bhalerao. Kernel Designs for Efficient Multiresolution Edge Detection and Orientation Estimation. *IEEE Trans. PAMI, March*, 1992.

[41] S. Levialdi. Basic Ideas for Image Segmentation. In O. D. Faugeras, editor, *Fundamentals in Computer Vision*, pages 239–261. Cambridge University Press, 1983.

[42] A. Montanvert, P. Meer, and A. Rosenfeld. Hierarchial Image Analysis Using Irregular Tessellations. *IEEE Trans. PAMI.*, 13:307–315, 1991.

[43] R. Wilson, S. C. Clippingdale, and A. H. Bhalerao. Robust Estimation of Local Orientations in Images Using a Multiresolution Approach. In *Proc. 5rd SPIE Conf. Vis. Comm. Image Proc.*, Lausanne, 1990.

[44] H. Knutsson. *Filtering and Reconstruction in Image Processing*. PhD thesis, University of Linköping, Sweden, 1982.

[45] S. Clippingdale. *Multiresolution Image Modelling and Estimation*. PhD thesis, Department of Computer Science, The University of Warwick, UK, September 1988.

[46] A. H. Bhalerao. *Multiresolution Image Segmentation*. PhD thesis, Department of Computer Science, The University of Warwick, UK, October 1991.

[47] G. Welch and G. Bishop. An Introduction to the Kalman Filter. Technical Report TR95-041, Department of Computer Science, Chapel Hill, NC, USA, 1995.

[48] M. Heath, S. Sarkar, T. Sanocki, and K. W. Bowyer. A Robust Visual Method for Assessing the Relative Performance of Edge-Detection Algorithms. *IEEE Trans. PAMI.*, 19(12):1338–1359, 1997.

[49] T-I. Hsu and R. Wilson. A Two-Component Model of Texture for Analysis and Synthesis. *IEEE Trans. Image Proc.*, 10(7):1466–1476, 1998.

[50] C-F. Westin, A. Bhalerao, H. Knutsson, and R. Kikinis. Using local 3d structure for segmentation of bone from computer tomography images. In *Proc. of Computer Vision and Pattern Recognition '97*, Puerto Rico, 1997.