

Feature Detection and Fusion for Intelligent Compression

P. G. Ducksbury[‡], M.J.Varga

Defence Evaluation & Research Agency, St Andrews Road, Malvern, Worcs, WR14 3PS, UK.

ABSTRACT

In previous work [1] a novel approach was described which used automatic target detection together with compression techniques to achieve intelligent compression by exploiting knowledge of the image content. In this paper an extension to this work is presented in which a set of standard feature detectors such as HV-quadtrees, approximate entropy and phase congruency are used as target discriminators. These detectors all attempt to find potential areas of interest within an image but will undoubtedly be slightly different in their estimates. A probabilistic (*Bayesian belief*) network is then used to fuse this information into a single hypothesis of 'interesting areas' within an image. A wavelet-based decomposition can then be applied to the image in which selective destruction of wavelet coefficients is performed outside the cued areas of interest (*in effect concentrating the wavelet information into the required areas*) prior to the encoding with a version of the progressive SPIHT encoder [8]. One of the difficulties with this approach can be when large quantities of wavelet coefficients are discarded, this can potentially lead to abrupt changes at a mask boundary resulting in (*visually*) undesirable effects in the reconstructed image. An improvement to this is to modify the fused feature image using morphology in order to arrive at a multi-level fuzzy mask. This can then be used to gradually reduce the significance of coefficients as the distance from the mask increases. Results will illustrate how this approach can be used for the detection and compression of airborne reconnaissance imagery

Keywords : Target Recognition, Intelligent Compression, Feature detection, Image Processing,

1. INTRODUCTION

The application of this work has been mainly in the area of analysing 8-bit airborne imagery. It is important that we can automatically detect targets of interest and be able to achieve good compression ratios as well as making sure that the important targets and areas of interest are not removed or degraded. As a result of these requirements coupled with the desire not to invent new and specific feature detectors, [1] described the use of quadtrees [2] and entropy [3-6] as methods for locating areas of interest within imagery. This has now been extended to also include results from phase congruency [27,28] as well as the combination of all three techniques using a belief network and the introduction of fuzzy masks.

A standard quadtree [2] based approach is modified so that it can be used as a target discriminator using the more advanced horizontal-vertical (HV) partitioning scheme. This partitioning attempts to make edges within the image run diagonally through a partition and thus divide a region into two subregions, if this is not possible it reverts to the standard decomposition into four sub-regions of equal size. Results will illustrate how the technique is more flexible than the standard quadtree resulting in significantly less regions in the decomposition. Entropy, approximate entropy and phase congruency have been used as an alternative to quadtrees and can lead to improved region cueing. In particular the phase congruency has proven to be an excellent technique for the detection of fine scale linear features. All techniques have their advantages and disadvantages for this reason a belief network was used to combine results from all three probabilistically.

The cued regions of interest obtained from the feature detection stage are subsequently compressed using an 'intelligent' image compression technique. The technique allows for selective compression to preserve detail in the key areas within the imagery using a wavelet based approach over the entire image but with different levels of compression at different image locations, we are hence cueing the wavelet information into desired areas. Various standard compression/coding algorithms where considered including MTWC [7], SPIHT [8,9] which is based upon an extension to the embedded zerotree wavelet

[‡] Correspondence : P.G.Ducksbury, Email : ducksbury@signal.dera.gov.uk

(EZW) work of Shapiro [10], CREW [11], Vector Quantisation schemes [12,13], Fractal [14], JPEG, wavelet, as well as the lossless JPEG (JPEG-LS) [15-17]. Of these the SPIHT algorithm was chosen as being the most suitable. Sections 2-4 describe the target detection using quadrees, entropy & approximate entropy and phase congruency whilst section 5 overviews the belief network approach to evidence fusion. Section 6 discusses the introduction of fuzzy masks whilst sections 7 and 8 deal with the actual compression and encoding aspects.

2. FEATURE DETECTION - QUADTREES

Quadrees are used as an attempt to do a generic and automatic cueing of areas of interest without the need to invent a new set of feature detection algorithms. A quadtree essentially a recursive decomposition of an image in which at each level the image is decomposed in its simplest form by a factor of 2 in both its x and y directions until some suitable criteria is satisfied. Here a textural measure based on edge data and statistical variance within an image portion is used.

More advanced partitioning methods are available such as HV partitioning which is more flexible than the standard quadtree as the position of the partition is not fixed where the image can be partitioned either horizontally or vertically and not by a direct subdivision by 2. The HV decomposition technique attempts to make image edges run diagonally through a partition. In addition there is also triangular partitioning in which the triangles can have any orientation. In the example shown in figure 1 an image with the standard and an improved (HV) quadtree overlayed is shown, whilst figure 2 shows part of the actual quadtree¹ for the standard version on the left of figure 1. Results have suggested that the HV technique can potentially offer a decomposition with up to 75% less rectangles. The HV version here is a modified one in which the algorithm will revert to the standard quadtree if no horizontal/vertical division can be easily detected. Wherever possible an image is split into two regions (*either horizontal or vertical*) using an analysis of the two profiles obtained by forming a summation of the current regions gradient information in the appropriate direction. The most dominant profile will ultimately determine the type of decomposition. Initial results show that the technique is feasible. In fact it can be seen from figure 1 that the quadtree version on the right is somewhat superior in terms of the reduced number of rectangles required to represent the area of interest. Although it will be noticed that both versions have failed to isolate all areas of interest but do provide an approximate cueing for an area.

3. FEATURE DETECTION - ENTROPY

Entropy is a measure of randomness (*or uncertainty*) within some given dataset, the more random the data the more entropy [4,19]. Entropy is defined as

$$H(X) \equiv \sum_{x \in W} p_x \log p_x$$

For example a region with a single gray level value would have a distribution p with just a single peak and an entropy measure of zero. For a section of sky the gray level distribution would be unimodal and therefore have a low entropy measure, whereas for a section with a more typical and varied image content the distribution would be more widespread and have greater entropy. The initial intention was to use the basic entropy as a measure of saliency within an image to highlight areas of interest. Unfortunately in the experiments carried out the standard entropy was very good at excluding areas (*e.g.* sky) but had a tendency to include too much other information within an image and any subsequent mask generated from this would be swamped by potential areas of interest.

Approximate Entropy [4,5,6] is a measure of unpredictability of a sequence of values.

$$f_m = -\frac{1}{N_m} \sum_{i=0}^{N_m} \log C_m^i \quad \text{and} \quad ApEn = f_{m+1} - f_m$$

Where N_m is the number of strings of length m which can be taken from a sequence and C_m^i is the number of strings which match string i .

¹ The tree shown here is not fully decomposed, the nodes indicated in black would be decomposed further. Branches without nodes are leaf nodes with no further decomposition, ordering for the quadtree is top left, top right, bottom left, bottom right.

Approximate entropy is similar to the standard entropy measure except that it is based on the frequencies of spatial configurations of pixels, rather than simply on the relative frequencies of individual pixel gray levels.

Figure 3 (*left*) shows the result of applying the approximate entropy to the image of figure 1. The darker areas indicate salient regions within the image and clearly show the main target of interest together with a certain (*but not excessive*) amount of background features.

4. FEATURE DETECTION - PHASE CONGRUENCY

Until recently, feature detection has largely comprised edge and corner detection. Edges are usually considered to be intensity changes that are locally one-dimensional whilst corners are intensity changes that are locally two-dimensional. The designs of most edge detectors have essentially been optimised for the detection and localisation of step changes in intensity. It is however well known that phase carries (*most of*) the important information about a signal [34], most current edge detectors developed for image analysis retain only amplitude information. Kovessi [27] points out, this not only causes them to fail to detect features of interest, it reduces markedly their invariance to intensity contrast, and makes them sensitive to thresholds.

Many intensity changes of interest in our applications are not at all step-like. There is a single characteristic that (*all*) image features (*luminance profiles that humans perceive as places of interest*) have in common, namely that in the (*short term, or windowed* [26]) Fourier domain (*all*) frequency components over a wide range of octaves are maximally in phase [35]. The actual angle at which this phase-congruency occurs is characteristic of the type of feature. eg for a +ve step in intensity all local (*windowed*) Fourier components will have a phase of 0, a -ve step will have phase π , a +ve ridge $\pi/2$, a -ve ridge $3\pi/2$. The observation suggests that the concept of a 'feature' is defined as an image location at which there is a local congruence of phase. The technique is invariant to local image contrast. That is, will generate the same features for compression irrespective of the contrast and brightness of the image. Phase Congruency [35] can be defined as

$$PC(x) = \max_{f(x) \in [0, 2\pi]} \frac{\sum_n A_n \cos(f_n(x) - \bar{f}(x))}{\sum_n A_n}$$

where A_n is the amplitude of the n^{th} Fourier component and $f_n(x)$ is the local phase of the Fourier component at x . The value $\bar{f}(x)$ that maximises the equation is the amplitude weighted mean local phase angle of all Fourier terms at the point being considered. However an alternative [36] is to search for peaks in the local energy function

$$E(x) = \sqrt{F^2(x) + H^2(x)}$$

where $F(x)$ is the signal with the DC component removed and $H(x)$ is the Hilbert transform of $F(x)$. For two dimensional imagery the one dimensional process is computed at a set of different orientations, refer Kovessi [27] for full details.

Figure 3 (*right*) illustrates this technique, standard edge detection techniques such as Sobel and Canny will perform very poorly on fine scale structure present in the images.

5. FUSION - BELIEF NETWORK

Previous work has considered the use of belief networks in image processing for the detection of urban regions, driveable regions in autonomous land vehicle imagery and improved vehicle detection in airborne imagery [29 - 32]. Pearl's Bayes networks are directed acyclic graphs, see figure 4 (*left*). In this graph nodes B, C and E represent different statistical information extracted from the image, whilst node A represents the '*belief*' in detecting an urban patch. A graph G is a pair of sets (V, A) for which V is non-empty. The elements of V are vertices (*nodes*) and the elements of A are pairs (x, y) called

arcs (links) with $x \in V$ and $y \in V$. Consider the simple network that is shown in figure 4 (left). The equations for computing belief and propagation of information are given in the following sections².

5.1 Belief Equations

Consider the link from node B to A then the graph G consists of the two subgraphs G_{BA}^+ and G_{BA}^- . These two subgraphs contain the datasets D_{BA}^+ and D_{BA}^- respectively.

From figure 4 it can be observed that node A separates the two subgraphs $G_{BA}^+ \cup G_{CA}^+ \cup G_{EA}^+$ and G_{AF}^- . Given this fact we can write the equation:

$$P(D_{AF}^- | A_i, D_{BA}^+, D_{CA}^+, D_{EA}^+) = P(D_{AF}^- | A_i) \quad (1)$$

by using Bayes rule the belief in A_i can be defined as:

$$\begin{aligned} BEL(A_i) &= P(A_i | D_{BA}^+, D_{CA}^+, D_{EA}^+, D_{AF}^-) \\ &= a \cdot P(A_i | D_{BA}^+, D_{CA}^+, D_{EA}^+) \cdot P(D_{AF}^- | A_i) \\ &= a \cdot P(D_{AF}^- | A_i) \cdot \left[\sum_{j,k,l} P(A_i | B_j, C_k, E_l) \cdot P(B_j | D_{BA}^+) \cdot P(C_k | D_{CA}^+) \cdot P(E_l | D_{EA}^+) \right] \end{aligned} \quad (2)$$

where α is taken to be a normalising constant and i, j, k and l range of the number of variables in A, B, C and E respectively. It can be seen from figure 4 that this equation is computed using three types of information:

- Causal support π (from the incoming links).
- Diagnostic support λ (from the outgoing links).
- A fixed conditional probability matrix (which relates A with its immediate causes B, C and E).

The equations which form the above information are given as follows. Firstly the causal support equations:

$$p_A(B_j) = P(B_j | D_{BA}^+) \quad (3)$$

$$p_A(C_k) = P(C_k | D_{CA}^+) \quad (4)$$

$$p_A(E_l) = P(E_l | D_{EA}^+) \quad (5)$$

Secondly, the diagnostic support equation is given by:

$$l_F(A_i) = P(D_{AF}^- | A_i) \quad (6)$$

Finally, the conditional probability matrix is defined to be:

$$P(A|B, C, E) \quad (7)$$

² The equations are derived along similar lines to those derived by Pearl [32] where in his example node A has just two predecessors and two successors.

The belief equation can now be rewritten to obtain the belief at node A based on the observations at B, C and E, e.g. the belief that an urban region is detected:

$$BEL(A_i) = a \cdot l_F(A_i) \cdot \sum_{j,k,l} P(A_i|B_j, C_k, E_l) \cdot p_A(B_j) \cdot p_A(C_k) \cdot p_A(E_l) \quad (8)$$

The belief at nodes B, C and E can be obtained from the equations:

$$BEL(B_j) = a \cdot p_A(B_j) \cdot l_A(B_j) \quad (9)$$

$$BEL(C_k) = a \cdot p_A(C_k) \cdot l_A(C_k) \quad (10)$$

$$BEL(E_l) = a \cdot p_A(E_l) \cdot l_A(E_l) \quad (11)$$

In other words the belief is the resultant product of causal support information, diagnostic support information and prior knowledge. The propagation equations described below are iterated to support belief of a certain event.

5.2 Propagation Equations

The propagation equations for the network are derived as follows, firstly the diagnostic support. From a previous analogy with equation (6) it can be formulated as:

$$l_A(B_i) = P(D_{BA}^- | B_i) \quad (12)$$

by partitioning the D_{BA}^- into its component parts, namely A , D_{AF}^- , D_{CA}^+ , D_{EA}^+ we can obtain

$$l_A(B_i) = a \sum_{j,k} \left[p_A(C_j) \cdot p_A(E_k) \cdot \sum_l l_F(A_l) \cdot P(A_l | B_i, C_j, E_k) \right] \quad (13)$$

likewise for $l_A(C_j)$ and $l_A(E_k)$

$$l_A(C_j) = a \sum_{i,k} \left[p_A(B_i) \cdot p_A(E_k) \cdot \sum_l l_F(A_l) \cdot P(A_l | B_i, C_j, E_k) \right] \quad (14)$$

and

$$l_A(E_k) = a \sum_{i,j} \left[p_A(B_i) \cdot p_A(C_j) \cdot \sum_l l_F(A_l) \cdot P(A_l | B_i, C_j, E_k) \right] \quad (15)$$

5.3 Causal Equations

These are defined using a similar analogy as follows.

$$p_F(A_i) = P(A_i | D_{BA}^+, D_{CA}^+, D_{EA}^+) \quad (16)$$

and from this we then derive the equation

$$p_F(A_i) = a \left[\sum_{j,k,l} P(A_l | B_i, C_j, E_k) \cdot p_A(B_j) \cdot p_A(C_k) \cdot p_A(E_l) \right] \quad (17)$$

An important point to realise from the equations (13 -15 & 17) is the fact that they demonstrate that the parameters λ and π are orthogonal to each other, i.e. perturbation of one will not affect the other. Hence evidence propagates through a network and there is no reflection at boundaries.

The idea behind using the belief network approach has been to combine multiple sources of information (*evidence*) in a network to achieve a more mathematically sound result than could be achieved by just applying a logical operation to merge all the information (*evidence*) together. Here, two scales proved sufficient together with three states (*low, medium and high*) for the evidence variable that represented the required region, the actual network used being a simple causal tree, figure 4 (*right*). By applying this technique a combined mask can be obtained a typical outline of which is shown in figure 7.

6. FUZZY MASKS

One of the problems that has appeared with the cueing of wavelet information is when large quantities of wavelet coefficients are removed and that it occurs close to an edge of a computed feature mask. This can cause an unnatural and abrupt transition in the wavelet coefficients (*eg this occurs around the middle of the upward pointing gun barrel in figure 7*). One simple solution to this problem is to extend the mask and to apply some increasing coefficient reduction (*not removal*) around the edge of the mask.

This process is achieved by employing a repeated application of a morphological dilation operation using a simple circular 5x5 kernel. The resulting multilevel mask is then used in the wavelet encoder to reduce the significance of coefficients by 0%, 25%, 50% and 75% as the distance from mask to background increases. This achieves the desired effect and also due to the increased size of the mask focuses the removal more onto the background of the imagery (*eg rows 1 and 2 of table 1*). Another advantage to this approach is that by gradually reducing the magnitude of coefficients there are less bits requiring to be encoded and subsequently transmitted (*eg rows 2 and 3 of table 1*).

7. INTELLIGENT IMAGE COMPRESSION

The motivation for using quadrees, entropy and phase congruency was an attempt to do a generic and automatic cueing of areas of interest without the need to invent new feature detection algorithms. In [1] a number of standard compression algorithms were considered in which the wavelet approach was chosen as the most suitable as we can target the wavelets to key areas and therefore apply different compression ratios in different image. The standard wavelet decomposition [21,33] is used together with the simplest (*and most localised*) Daubechies-4 filter kernel [18]. Figure 5 illustrates this for a 3 level decomposition in which LH indicates Low-horizontal and High-vertical, HL : High-horizontal and Low-vertical, HH : High-horizontal and High-vertical whilst LL (*not shown*) is decomposed into the next level of the pyramid. The inverse wavelet transform then being used to reconstruct the original image.

Figure 6, shows a standard result from wavelet compression, namely that a large percentage of the wavelet coefficients in the wavelet scale space image can actually be removed yet still produce a visually good reconstruction. It is apparent therefore that a relatively small percentage of the wavelet coefficients are concentrated in the highly salient areas of the image. This is itself the main principle behind the technique used here, i.e. use a target mask generated by some saliency operator to prevent removal of coefficients in key areas of the image. When this is combined with a suitable encoder an efficient progressive target based compression system is produced.

8. ENCODER

The SPIHT (Set Partitioning In Hierarchical Trees) algorithm³ is a wavelet based compression scheme offering good image quality with high PSNR⁴ it is designed for progressive transmission producing a fully embedded coded file. SPIHT [8,9] is based upon an extension to the embedded zerotree wavelet (EZW) work of Shapiro [10]. It is a coding method so any artifacts produced may be due to the wavelet transform and not the coding process. The process differs from conventional wavelet compression in the way in which it encodes the wavelet coefficients. It can also be used for lossless compression as it codes the bits of the image wavelet transform coefficients in a bit-plane sequence, thus if all bits are encoded the image

³ The implementation of the SPIHT algorithm used here does not contain any arithmetic coding.

⁴ $psnr = 20 \log_{10} \left(\frac{2^b}{\sqrt{rmse}} \right)$ PSNR is the Peak Signal to Noise ratio and b is the bit depth with *rmse* being the root mean square error.

can be recovered perfectly⁵. As with other techniques error protection is important, but SPIHT's embedded coding scheme means that information is sorted according to importance hence the need for error protection reduces towards the end of the file. The technique is based on three ideas : a tree based representation of the wavelet coefficients (figure 5 - right), partial magnitude (*so that the most important information may be transmitted first*) ordering of the coefficients and an ordered bit plane transmission of refinement bits for the coefficients. As a result of this, both the encoder and decoder know the structure of the data and the algorithm therefore no explicit ordering information needs transmitting. Hence if a coefficient at a given level in the hierarchy is insignificant then those children of that node are also likely to be insignificant. The algorithm can also be efficiently implemented using a few list structures.

Figure 10 shows the progressive nature of the SPIHT algorithm for the combined mask⁶. The first column (top - bottom) represents 290:1, 130:1 and 60:1 compression whilst the second column (top - bottom) 40:1, 20:1 and 10:1 compression. For comparison figure 9 shows results of this with the standard JPEG at a ratio of 50:1 (0.16 bits per pixel) with the PSNR for these images being 26.91 and 22.69 respectively.

9. CONCLUSIONS

The improved mask generated by fusing feature detectors and introducing a fuzzy representation now more accurately represents features such as for example the gun barrel, which was partially cropped when using the quadtree generated mask in isolation. Table 1 shows the PSNR figures for several versions of the approach we have described, using a normal fused mask compared with several version of the fuzzy mask. Although the simple fused mask appears better with a psnr of 33.92 this is due to the additional information present in the image and therefore additional information that requires encoding and transmission. These results illustrate the fact that psnr is probably not the best measure to be used for comparing results for the more content based (*intelligent*) compression algorithms. All three sets of results produced reconstructed imagery for which it would have been difficult for the human observer to notice any differences.

There are a number of issues worth mention in particular, although object detection is a difficult task it is still worthwhile to produce object masks. If we remove a certain amount of information (*e.g. wavelet coefficients*) then it is more intuitive to remove them from undesirable and less salient parts of an image rather than removing them indiscriminately⁷. Despite the fact that the use of a state-of-the-art codec is being employed it is still advantageous to attempt to remove non essential information prior to the encoding process. Typically in the examples shown here, to use a SPIHT codec directly without any masks and removal of coefficients would have required an additional ~ 500 Kbits to ~ 1 Mbits. Technique have been shown here that can locate salient parts of an image using several approaches, combine these into a cohesive mask and then use this to cue the wavelet information to obtain improved compression results.

Photographic interpreters have analysed these results by being presented with a set of the progressive full sized images shown in figure 10, starting with the most compression and moving towards the least. Their analysis being one of a hierarchical refinement as the image presented was improved. Therefore protecting key areas and thus enabling these to be restored more quickly does provide added value to the recognition process.

10. REFERENCES

1. Ducksbury P.G., 'Target detection and intelligent image compression', SPIE Aerosense 2000, April 24-28, 2000
2. Samet H., Tamminen M., 'Computing geometric properties of images represented by linear quadtrees', IEEE Trans PAMI, vol 7, no 2. March 85.
3. Gilles S., 'Robust description and marching of images', PhD Thesis, Department of Engineering Science, Oxford University, 1998
4. Pincus S.M., 'Approximate entropy : statistical properties and applications', Commun. Statistical Theory Meth., 21:3061-3077, 1992.

⁵ There are issues of precision here, as theoretically lossless compression is only possible if the coefficients are encoded using infinite precision. The remedy to this is to use an integer multiresolution transformation referred to as the S+P transform. This is claimed to be as efficient as most lossless encoders.

⁶ The original imagery was 512x512, 8 bit.

⁷ The work described here mirrors what is going on within JPEG-2000 - in particular the encoding of arbitrary shaped regions of interest using wavelet based compression although it approaches the problem from a different perspective. This part of the standard should be available later this year for evaluation.

5. Mortimer V., 'Content-based Image Compression', First year report, Department of Engineering Science, Oxford University, Oct 1999.
6. Rukhin A.L., 'Approximate entropy for testing randomness', Tech report, Department of Mathematics & Statistics, Univ. of Maryland, 1998.
7. Hounj-Jyh Wang et al. 'Multi-Threshold Wavelet Codec (MTWC)', ISO/IEC JTC1/SC29/WG1 N819, Convergence Phase Meeting, Geneva, March 98. <http://biron.usc.edu/~houngjyh/cerimon/demo/97-0005/index.htm>
8. Said A., Pearlman W.A., 'A new fast and efficient image codec based on set partitioning in hierarchical trees', IEE Trans on Circuits and Systems for video technology, Vol. 6, June 1996, <http://ipl.rpi.edu/SPIHT/spiht0.html>
9. Manduca A., Said A., 'Wavelet compression of medical images with set partitioning in hierarchical trees', SPIE Symposium on Medical Imaging, Cambridge, MA, March 96, <http://ipl.rpi.edu/SPIHT/spiht0.html>
10. Shapiro J.M., 'Embedded image coding using zerotrees of wavelets coefficients', IEEE Trans Signal Processing, vol. 41, pp 3445-3462, Dec 93.
11. Boliek M., et al., 'Next generation image compression and manipulation using CREW', ICIP-97, Santa Barbara, CA, Oct 97, <http://www.crc.ricoh.com/CREW/>
12. Rizvi S.A., Nasrabadi N.M., 'Predictive residual vector quantisation', IEEE Trans Image Processing, vol. 4, no 11, pp 1482-1495, Nov 95.
13. Oehler K.L., Cossman P.C., Gray R.M. May J., 'Classification using vector quantisation', In Proc 25th Annual Asilomar Conf on signals and computers, pages 439-445, Pacific Grove, CA, Nov 91
14. Fisher Y., 'Fractal Image Compression', SIGGRAPH 92 course notes, available via <http://inls.ucsd.edu/y/Fractals/>
15. <http://www.hpl.hp.com/loco/> - Lossless JPEG download page
16. Preuss D., 'Comparison of 2-D facsimile coding schemes', IEEE Press, 1976.
17. ISO Final Draft standard, 'Lossless and near-lossless coding of continuous tone still images, JPEG-LS', FCD14495, ISO/IEC JTC1/SC29
18. Daubechies I., 'Wavelets', S.I.A.M., Philadelphia, 1992.
19. MacKay D., 'A short course on Information Theory', Feb 1995, <http://131.111.48.24/pub/mackay/info-theory/course.html>
20. Fawcett R., 'Efficient and practical image compression', PhD thesis, Dept of Eng. Sci, Oxford University, Jan 97.
21. Clarke R.J., 'Digital compression of still images and video', Academic Press, 1995.
22. Held G., Data & Image Compression : Tools and Techniques, 4th edition, John Wiley 1996.
23. Schafer R., MPEG-4 : a multimedia compression standard for interactive applications and services, Electronics and Communication Engineering Journal, December 1998., pp 253-262.
24. Rabbani M., Jones P.W., 'Digital Image Compression Techniques', SPIE Tutorial texts in Optical Engineering, Vol. TT 7, 1991.
25. Xu X., 'An algorithmic study on lossless image compression', IEEE Data Compression conference DCC-96, Snowbird, Utah, April 96.
26. Mallat S., 'A wavelet tour of signal processing', Academic Press, 1998.
27. Kovesi P., 'Image features from Phase Congruency', Journal of computer vision, MIT Press. <http://www.cs.uwa.edu.au/~pk/Research/research.html>
28. Schenk V., 'Visual identification of fine surface incisions', Dept. Engineering Science, Oxford University, UK, 1st year report, July 1998.
29. Ducksbury P.G., 'Parallel texture region segmentation using a Pearl Bayes Network', British Machine Vision Conference, BMVC-93, University of Surrey, Guildford, UK, pp187-196, Sept 93.
30. Ducksbury P.G., Booth D.M., Radford C.J., 1995, 'Vehicle detection in infrared linescan imagery using Belief Networks', IEE IPA-95, 415-419
31. Varga M.J., Ducksbury P.G., 'Infra-Red Thermography : Techniques and Applications', Handbook of Pattern Recognition and Computer Vision (2nd Ed), Eds. Chen, Pau & Wang, World Scientific, pp 891-924, 1999.
32. Pearl J., 1986, 'Fusion, Propagation and structuring in Belief Networks', AI, 29, 241-288
33. Castleman K.R., 'Digital Image Processing', Prentice Hall 1996.
34. Morrone M.C., Burr D.C., 'Feature detection in human vision : A phase-dependent energy model', *Proc Royal Society*, London, B, 235:221-245, 1988.
35. Morrone M.C., Owens R.A., 'Feature detection from local energy', *Pattern Recognition Letters*, 6:303-313, 1987.
36. Venkatesh S., Owens R.A., 'An energy feature detection scheme', Int Conf on Image Processing, pp 553-557, 1989, Singapore.
37. Christopoulos C., Askelof J., Larsson M., 'Efficient methods for encoding regions of interest in the upcoming JPEG2000 still image coding standard', IEEE Signal processing letters, vol. 7, no 8, pp 247-249, Sept 2000.

38. Creusere C.D., Nevel A.V., 'Automatic target recognition directed image compression', Journal of Aircraft, vol. 36, no 4, pp 626-631, July-August 1999.

(c) British Crown Copyright 2001/DERA,
Published with the permission of the controller of Her Britannic Majesty's Stationary Office.
This work was carried out as part of the Technology Group (TG10) of the MoD Corporate Research Programme

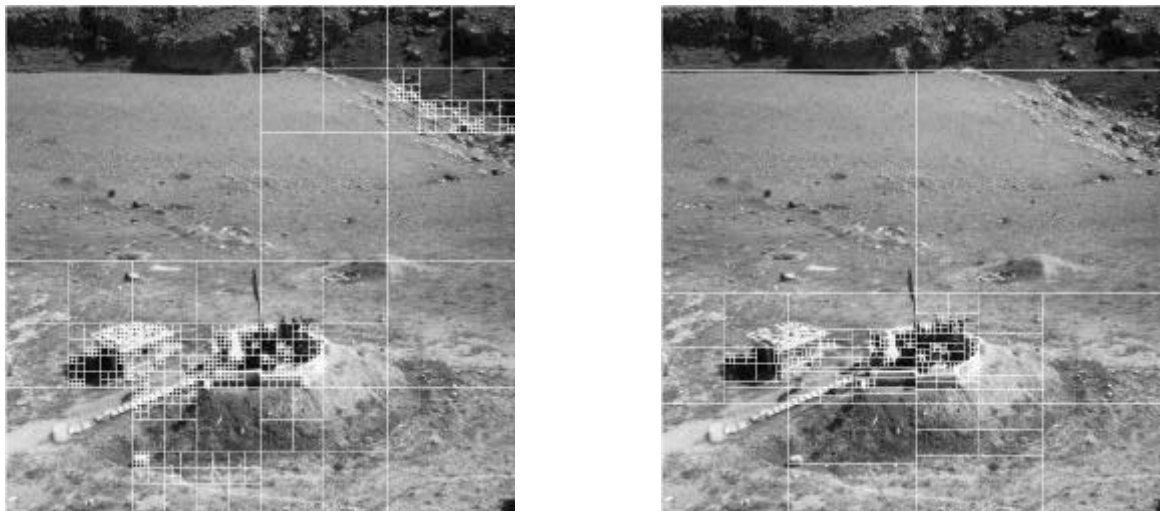


Figure 1, Left - Standard quadtree decomposition, Right - H-V decomposition

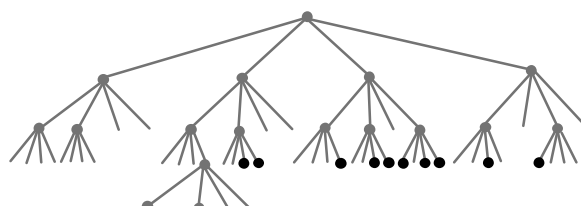


Figure 2, Initial layers of standard quadtree

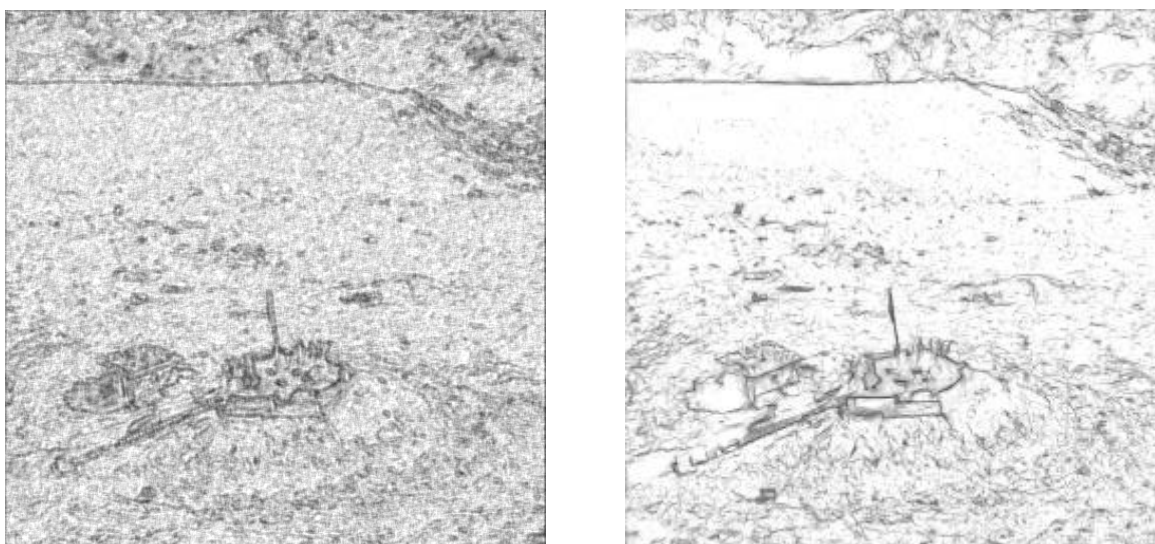


Figure 3, Left – Approximate Entropy, Right - Phase Congruency

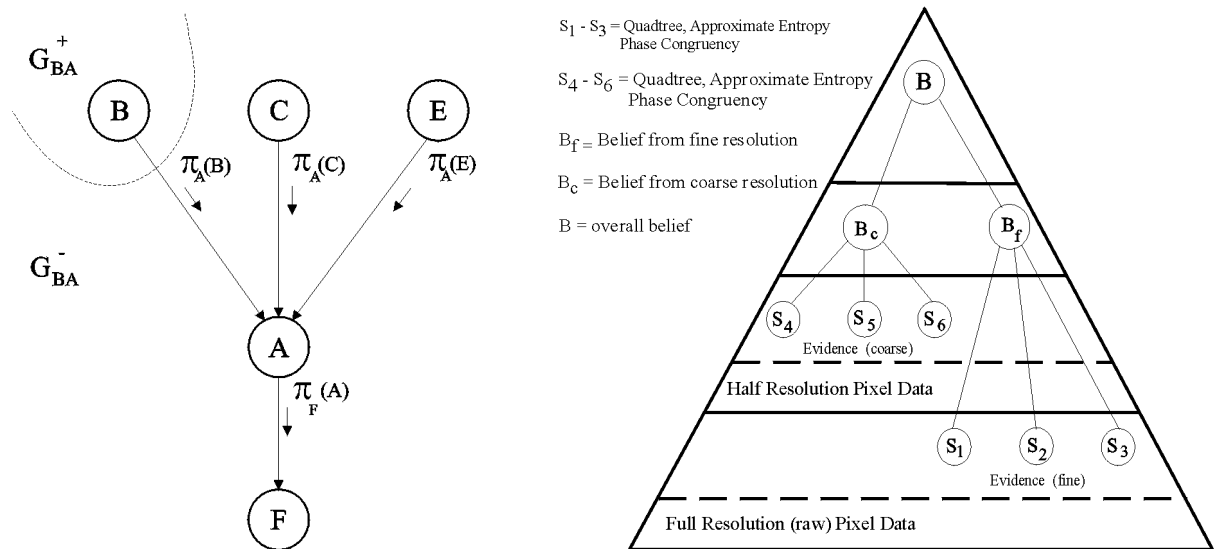


Figure 4, Belief Network

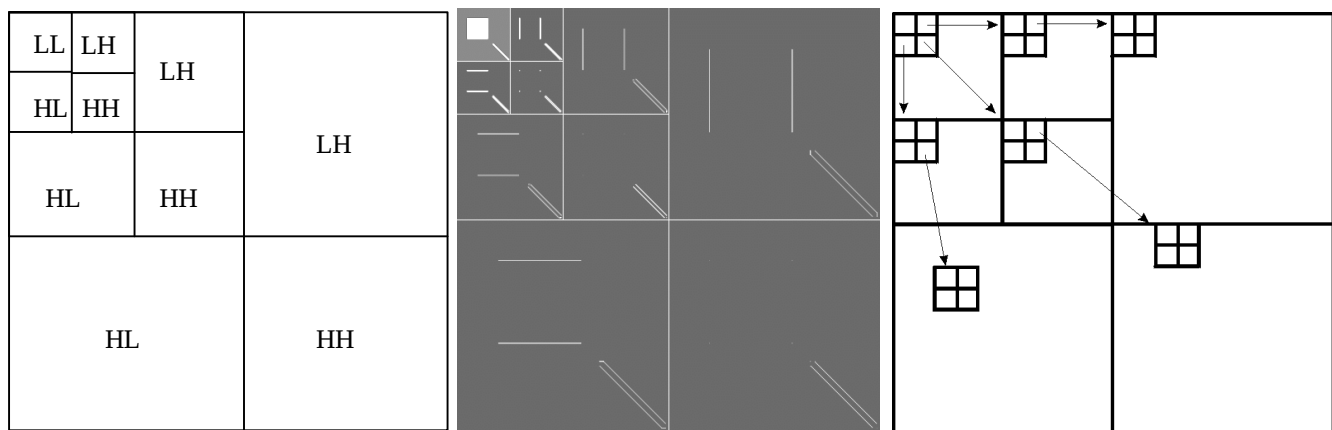


Figure 5, Wavelet Decomposition



Figure 6, Wavelet reconstruction after removal of 30%, 60% and 95% of weakest coefficients.



Figure 7, Outline of combined mask.

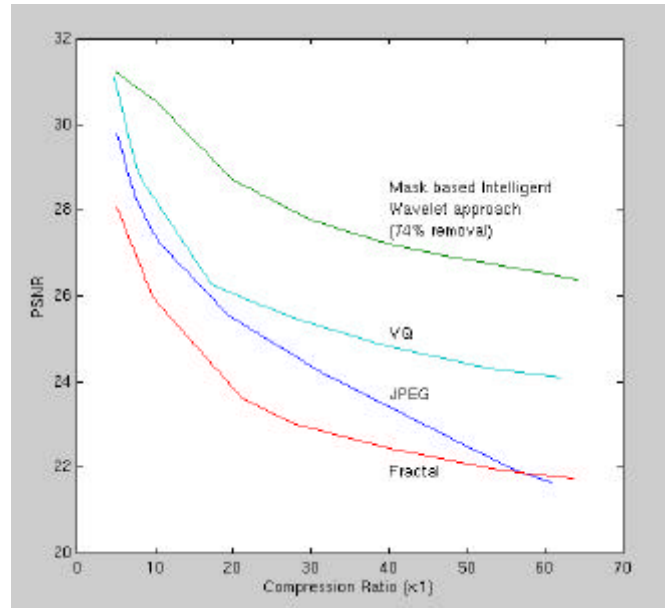


Figure 8, PSNR verses compression ratio



Figure 9, Target Detection and Compression @ 50:1 (0.16 bits/pixel). Left - Combined Fuzzy Mask-based wavelet, Right - JPEG for comparison. The PSNR measurements being 26.91 and 22.69 respectively.

Object Mask Type	Total Bits (to encode entire image)	PSNR (full Reconstruction)
Normal (fused) mask	938531	33.92
Fuzzy (fused) mask - no coefficient reduction	735187	31.58
Fuzzy (fused) mask - with coefficient reduction	699347	31.37

Table 1, Mask type vs. file size and PSNR

PSNR
22.99



PSNR
27.21



PSNR
25.00



PSNR
28.73



PSNR
26.53



PSNR
30.55



Figure 10, Progressive Target Based Reconstruction - I (Combined fuzzy mask & 74% removal in background)
Left Column, Top – Bottom 290, 130, and 60:1, Right Column, Top - Bottom 40, 20 and 10:1 compression.