Voronoi seeded colour image segmentation

Abstract

The goal of the segmentation scheme presented is to combine **edge** *and* **region** *information to achieve a stable segmentation. The segmentation scheme presented is designed to operate on general home and stock photographs, it returns a comprehensive region-based description of the visual content of an image (including a distinction between smooth and textured regions and a description of the internal properties of the later).*

A **colour edge detector** *is presented, where hue difference is weighted more heavily than brightness difference. Seed points for region growing are derived from the colour edge image as the peaks in the associated* **Voronoi image***. Regions are grown using gates on pixel colour relative to region central colour and region edge pixel colour. This permits regions to encompass shading gradients. Image edges act as hard barriers during region growing. A* **discrete feature based texture model** *is derived and then used to unify groups of smaller regions into extended textured regions.*

The segmentation scheme is designed to facilitate image retrieval and has been tested on a corpus of over 40000 images and has been found to be robust.

Keywords*: image segmentation, colour edge detection, textured region properties, region based image retrieval.*

1 Introduction

Image segmentation remains a problem. Qualitatively segmentation is the process of grouping together pixels which are semantically linked. This paper takes the view that in order to achieve a meaningful segmentation describable low level features must be extracted first and subsequently linked together using a series of opportunistic grouping algorithms. At the lowest level the only information available is local similarity. The driving motivation behind the propounded segmentation scheme is then that *edges can be found* and *homogeneous regions can be found* and then the two can be unified to give a very convincing segmentation.

Existing segmentation schemes are largely either texture-based or edge-based and are designed to operate on grey-scale images. Huttenlocher [6] uses a graph splitting approach with intensity variation between neighbouring pixels providing weighting between nodes. Shi [16] demonstrates a grouping mechanism based on normalised cuts but the proposed algorithm remains slow, requiring the manipulation of large sparse matrices. It is also worth noting that the method given in [16] takes account of a single relation type between pixels, essentially reducing the algorithm to hierarchical colour clustering.

Other graph based approaches exist [8]. Texture based approaches include Puzicha [12] which uses Gabor filters to generate texture features together with a pairwise co-occurrence model, and Forsyth [7] which distinguishes textured and un-textured areas and groups regions of repeated structure.

Direct edge-based segmentation is also possible [4, 14, 13, 2], but it is difficult to establish correspondence between image edge features and object or region boundaries. Supporting information from the areas between edges greatly improves the robustness of a segmentation.

The view is taken that all image boundaries are significant and that the smallest scale of detail present in the image should be available to the routines assembling a description of an image. Small features explicitly exist and should not be immediately subsumed by over-crude segmentation.

This paper is laid out as follows. A full **three colour edge detector** is presented (if applied to a grey scale image it defaults to a *Canny* type edge detector). Seed points for region growing are found as the peaks in the Voronoi image computed from the output of the colour edge detector. Regions are grown using gates on pixel colour relative to region central colour and region edge pixel colour. This permits regions to encompass shading gradients. Image edges act as hard barriers. Example results of the segmentation scheme are given.

A texture model is also given. The model functions by finding extended *ridges*, breaking the extended ridge map into short discrete chunks, grouping these chunks according to colour and orientation and then justifying these groups with the underlying Voronoi segmented regions.

A suggestion is made to use the Corel image library 1 as a standard test collection for comparing segmentation algorithms.

2 Colour edge detection

In this model for colour edge detection change in brightness and change in colour are weighted differently. If $R(i,j)$, $G(i,j)$, $B(i,j)$ represent the red, green and blue values at a pixel then the change in brightness $I(i,j)$ in the i direction is given by,

$$dR_i(i,j) = R(i-2,j) + \frac{R(i-1,j)}{2} - \frac{R(i+1,j)}{2} - R(i+2,j)$$

$$dI_i = dR_i + dG_i + dB_i$$

and with magnitude of change in colour given by

$$\begin{aligned} dC &= \sqrt{(}(dB_i - dG_i)^2 + (dR_i - dB_i)^2 + (dG_i - dR_i)^2 \\ &+ (dB_j - dG_j)^2 + (dR_j - dB_j)^2 + (dG_j - dR_j)^2). \end{aligned}$$

The weighted total change dT (dropping the indices (i,j)) is

$$dT = dI_i^2 + dI_j^2 + 3.0dC$$

The choice of 3 as the weighting factor in favour of colour change over brightness change is empirical but has been found to be effective. Local orientation (for use in the non-maximum suppression step of the edge detection process) is defined to be in the direction of the maximum colour gradient. dT is then the edge strength fed to the non-max suppression and hysteresis edge-following steps (which follow *Canny* [4]). Figure 1 shows a coloor image and the output of the colour edge detector. It was not found necessary or desirable to introduce a smoothing step before taking the three pairs of directional derivatives.

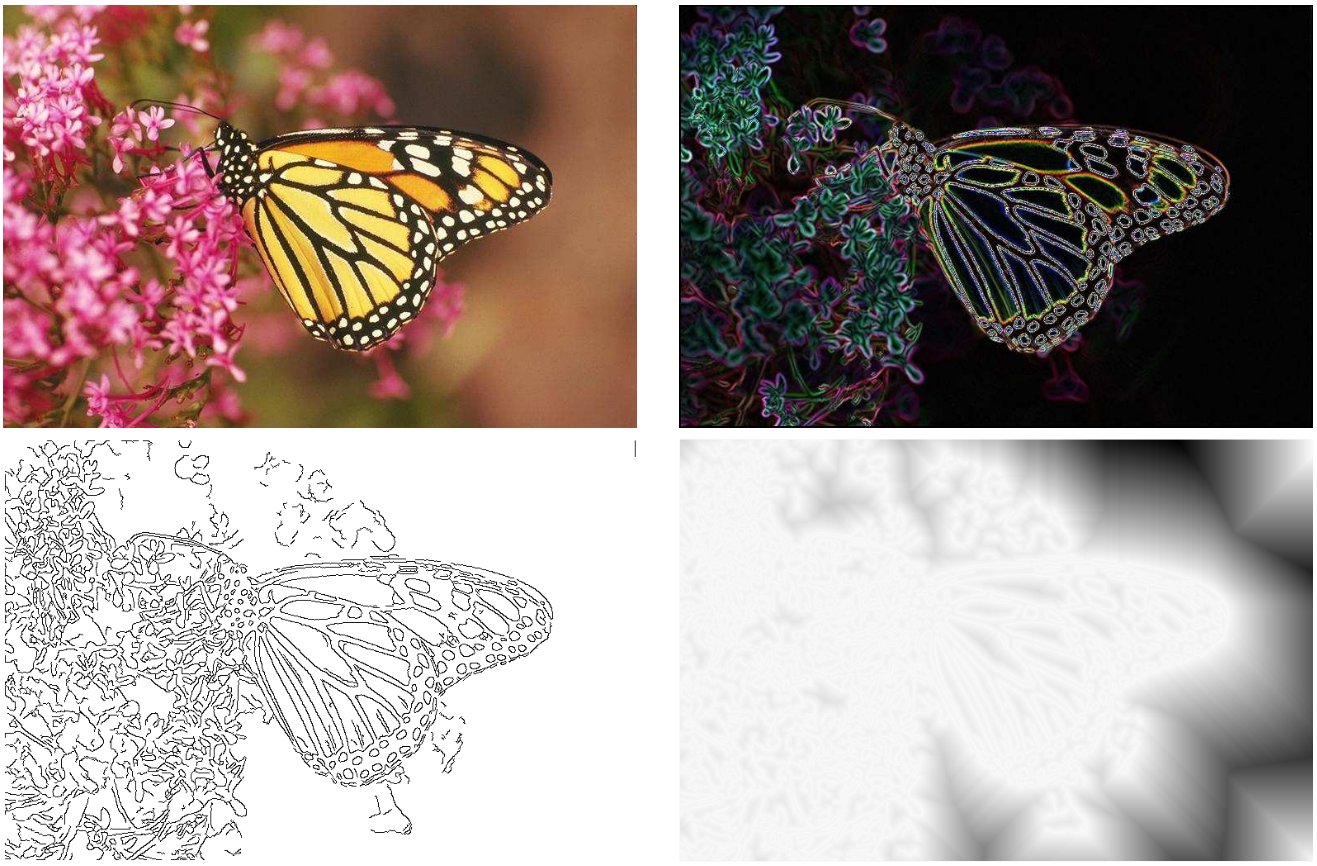

Figure 1: *Example colour image from the Corel stock image library.*
Full three colour derivative of the image in figure 1.
Edges derived using the three colour edge detector given above.
Voronoi image computed from the edge image to the left. Here the darker a pixel is the further it is from an edge.

3 Voronoi based region segmentation

3.1 Seed point generation

A *Voronoi* image is generated from the edge image. The intensity at each pixel in a Voronoi image is the distance to the closest edge. The peaks in the Voronoi image are then used as seed points for region growing. Figrue 1 shows the edge map derived from the colour edge detection process and the corresponding Voronoi image.

3.2 Smooth region growing

Regions are grown out from seed points. Unassigned pixels on the boundary of a growing region are assigned membership in that region if the difference in colour between the candidate and boundary member pixel is less than one threshold and the difference in colour between the candidate and the mean colour of the region is less than a second larger threshold. If a pixel is accepted into a region the mean colour of the region is then updated. Addition of a pixel to a region leads to additional candidate boundary pixels, these are added to the end of a candidate boundary pixel queue. Edges recovered from the image act as hard barriers through which regions are not allowed to grow.

Figure 2 shows the initial regions grown from a set of Voronoi seeds: black areas are either edges or have not yet been assigned region membership. A competitive morphological step is applied to absorb unassigned pixels into their best matching neighbouring region. Figure **??** shows the resulting segmentation after this morphological clean up steps have been applied to grow regions out to edges.

Figure 2: *Initial regions grown from Voronoi seeds with tight membership gates on new pixel colour. Secondly the gates of pixel colour are relaxed and regions grown out to images edges. Thirdly edges are competitively morphed into adjacent regions and similar colour regions are merged.*

Adjacent regions of similar colour are then merged and edges competitively merged into regions.

Figure 2 shows the coloured version of the final segmentation.

3.3 Voronoi segmentation results

Assessing segmentation results is problematic. For all but trivially simple examples (e.g. simulated blocks world) ground truth is not available. There is no accepted metric to compare one segmentation scheme with another and there is no public domain set of segmented images which may be used to provide a qualitative bench mark. The divide between *small regions* and *texture* remains almost entirely subjective.

The nearest thing available to a reference set of images is the Corel Stock Photo Library number 1. Figures 3 to 7 show pairs of images and their Voronoi segmented versions. These images are taken from the free sampler CD from the above collection. Segmented versions of all 120 images from this CD (at size 512 by 768 pixels) can be found at [3].

Subjectively, the Voronoi segmentation results on the Corel images look quite convincing in terms of capturing the essential features of the images in a relatively compact form.

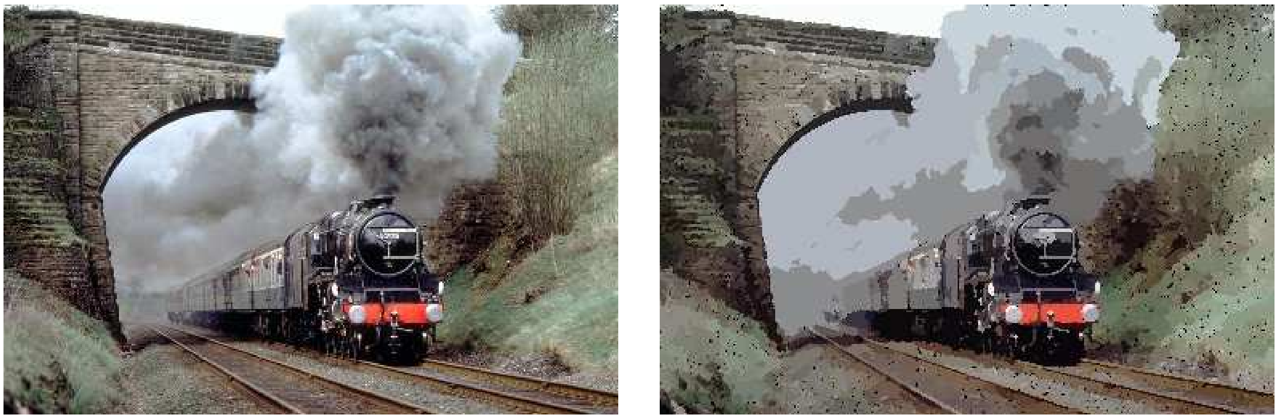

Figure 3: *Image 009 from Corel's free sampler CD and its corresponding Voronoi segmentation.*

Figure 4: *Image 010 from Corel's free sampler CD and its corresponding Voronoi segmentation.*

The above Voronoi based segmentation scheme performs well on images with large smooth regions. Heavily textured regions tend to be fragmented. The following discrete feature based

Figure 5: *Image 020 from Corel's free sampler CD and its corresponding segmentation.*

Figure 6: *Image 040 from Corel's free sampler CD and its corresponding segmentation.*

texture model provides a robust means of unifying fragmented regions into larger textured regions.

4 A discrete feature based texture model

The debate in the computer vision community over what comprises texture remains unsettled. The *texton* model [10] of texture perception was abandoned due to practical difficulties in defining and finding satisfactory low-level features. Probabilistic representations of local appearance have been tried [5, 18, 15], however such approaches are typically slow and suffer from only being able to model limited scope of interaction (either spatial extent or range of filter response). It could be argued that the nature of the interaction of between texture elements changes through the range of scale over which a texture is visible making multi scale approaches attractive [11, 17].

The texture model used in this paper functions as follows. Extended connected ridges are found (figure 8). A pixel is defined to be a ridge pixel if the magnitude of the second derivative operator applied to a grey-scale version of the original image exceeds a threshold. The ridge network is then broken up into compact 30 pixel groups or *features*. Figure 8 shows an image and its ridge network broken up into small discrete compact texture features. The orientation of each feature is computed from its second moment (about its center of mass). The colour of the

Figure 7: *Image 095 from Corel's free sampler CD and its corresponding segmentation.*

feature is computed from its footprint in the original image, and it connectivity to neighbouring features noted.

Net-like structures may be recovered from images as a by product of the texture model (figure 9). The topology of connected ridge maps is analysed for *holes* (singly connected enclosed regions of not ridge), if a ridge map has more than five *holes* is is deemed to be a net.

Oriented texture features are clustered using a greedy, sequential, agglomerative clustering algorithm. The set of pairwise distances between proximate *features* is computed. The distance between a pair of *features* is in this case the Euclidean distance between their RGB colours. The list of pairwise distances is ordered by increasing distance. Each pairwise relation is used in turn either to start a texture cluster center (here referred to as a *clique*, or to add a feature to a clique or to amalgamate two existing cliques. The result of this clustering process is shown in figure 10.

Cliques of features are then justified with the underlying Voronoi region segmentation. If the total footprint of texture features in a region exceeds a fixed percentage of the area of that region then the region is assigned membership in the corresponding dominant *clique*. This gives a set of regions found to be textured (as opposed to being smooth) and a set of *texture features* associated with each group of regions sharing a common dominant texture. Figure 10 shows a Voronoi segmented image its texture feature cliques and the recovered textured regions.

Internal texture structure of a textured region can then be described by an orientation histogram the *features* and a **Pairwise Geometric Histogram** (**PGH**) [1, 9] of relative orientation verses mutual separation. Contractions of the PGH are possible and desirable for image description for retrieval tasks but discussion of these is beyond the scope of this paper. The colour of a textured region then has two components, the mean colour of the texture features and the mean colour of the rest of the region.

The segmentation scheme has been applied to all 40000 images from Corel image library 1 and 2. Anecdotally the segmentation results look very good.

5 A note on comparing segmentations

It would be desirable if segmentation schemes could be compared is some more realistic way than eyeballing chosen examples given in a paper. This requires agreement on a large corpus of test images. A case could be made for using the (royalty free) Corel image library 1. 16 bit

Figure 8: *Image 030 from Corel's free sampler CD and its ridge map broken into discrete texture features.*

Sun raster files would seem a reasonable choice for representing region maps. Inevitably debate will exist on where the divide between low level feature and textured region should be draw and an arbitrary human operator decision may be necessary to provide an initial segmentation benchmark.

6 Conclusions

This paper presents a robust and competent segmentation scheme designed to operate on generic photographs. The segmentation scheme returns smooth and textured regions together with colour descriptors and internal texture structure descriptors. The segmentation scheme has been applied to a test corpus comprising 40000 images from Corel image library 1 and 2. The segmentation results look very good although this statement cannot currently be quantitatively justified.

7 Acknowledgements

The authors would like to acknowledge directional guidance and encouragement from Prof Andy Hopper, the members of the DART project at AT&T Laboratories Cambridge, and Kerry Rodden

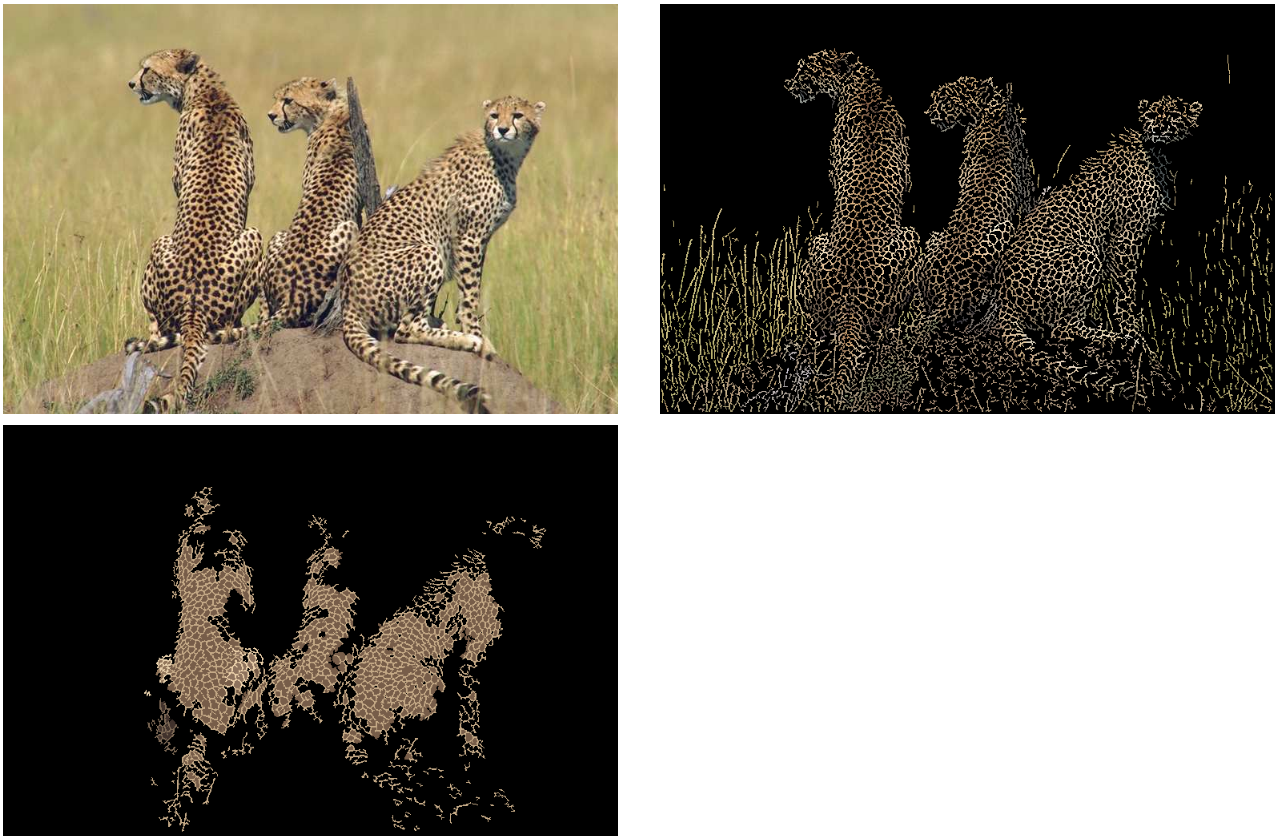

Figure 9: *Image 042 from Corel's free sampler CD, its ridge map broken into discrete texture features and the connected net structures extracted from the ridge map.*

of Cambridge University Computer Lab.

References

[1] A. Ashbrook, N. Thacker, and P. Rockett. Pairwise geometric histograms : A scaleable solution for the recognition of 2d rigid shape. Technical Report 94/30, Sheffield University, Electronic Systems Group, 1994.

[2] I. Biederman. Matching image edges to object memory. In *Int. Conf. on Computer Vision*, pages 384–392, 1987.

[3] AT&T Laboratories Cambridge. Sample image segmentations. In `http://www.uk.research.att.com/~das/iccv99sampler.html`.

[4] J.F. Canny. Finding edges and lines in images. Master's thesis, MIT, Cambridge, USA, 1983.

[5] I. M. Elfadel and R. W. Picard. Gibbs random fields, cooccurrences and texture modeling. *PAMI*, 16(1):24–37, 1994.

[6] P. Felzenswalb and D. Huttenlocher. Image segmentation using local variation. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 98–104, 1998.

[7] D. Forsyth, J. Malik, M. Fleck, and J. Ponce. Primitives, perceptual organisation and object recognition. Technical Report http://HTTP.CS.Berkeley.EDU/ daf/vr11.ps.Z, University of California, Berkeley, Computer Science Division, 1997.

[8] Y. Gdalyahu, D. Weinshall, and M. Werman. A randonized algorithm for pairwize clustering. Technical Report ?, The Hebrew University, Jerusalem, 1998.

[9] B. Huet and E. Hancock. Fuzzy relational distance for large-scale object recognition. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 138–143, 1998.

[10] B. Julesz and J.R. Bergen. Textons, the fundamental elements in preattentive vision and the perception of textures. *Bell Systems Technical Journal*, 62:1619–1644, 1983.

[11] S. Peleg, J. Naor, R. Hartley, and D. Avnir. Multiple resolution texture anlysis and classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:518–523, 1984.

[12] P. Puzicha, T. Hofmann, and J. Buhmann. Non-parametric similarity measures for unsupervised texture segmentation and image retrieval. In *IEEE International Conf. on Image Processing*, pages 267–272, 1997.

[13] P. Rosin. Edges: Saliency measures and automatic thresholding. Technical note no. i.95.58, Ispra, 1995.

[14] C. Rothwell, J. Mundy, B. Hoffman, and Nguyen V.D. Driving vision by topology. Technical Report 2444, INRIA, 1994.

[15] A. Sarkar, K. Sharma, and R. Sonak. A new approach for subset 2-d ar model identification for describing textures. *IEEE Transactions on Image Processing.*, 6(3):407–413, 1997.

[16] J. Shi and J Malik. Normalized cuts and image segmentation. In *IEEE Conf. Computer Vision and Pattern Recognition*, 1997.

[17] Z. Y. Xie. *Multi-scale analysis and texture segmentation.* PhD thesis, Oxford University, 1994.

[18] S. Zhu, Y. Wu, and Mumford D. Frame: filters, random fields, and maximum entropy principle. Technical Report TR 95-02, Harvard Robotics Lab, 1995.

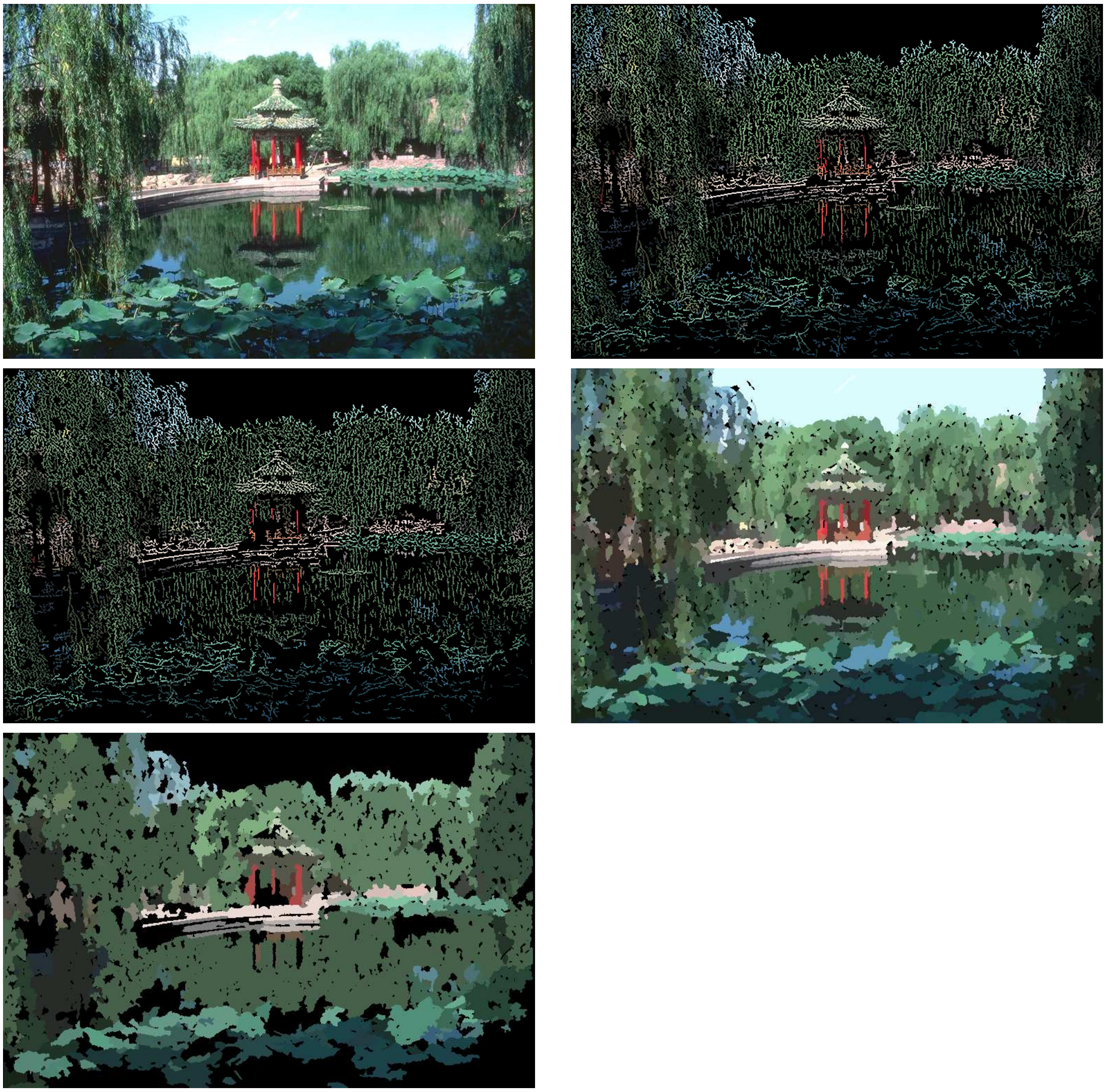

Figure 10: *Image 076 from Corel's free sampler CD, its ridge map broken into discrete texture features, cliques of clustered texture features, Voronoi segmented version and recovered textured regions.*