Multi-Level Probabilistic Relaxation

M. Petrou, M. Mirmehdi, and M. Coors*
Centre for Vision, Speech, and Signal Processing
University of Surrey
Guildford GU2 5XH, United Kingdom
email: M.Mirmehdi@ee.surrey.ac.uk

Abstract

We present a multi-level probabilistic relaxation scheme appropriate for image segmentation on the basis of features computed at various resolution levels. At each level, we incorporate contextual information computed from the features in the coarser scale representation of the scene, as well as information from the features computed in the finer scale representation, until a full segmentation is achieved at the highest level of resolution. The features used can be textural, colour, or the results of any other type of measurements performed on the image. We demonstrate our approach by the segmentation of colour textures.

1 Introduction

Probabilistic relaxation is a powerful technique that allows the incorporation of contextual information in labelling problems. The theoretical foundations of the approach have been developed in a series of papers by various researchers (e.g. [9, 7, 10, 4, 1]), and the method has been applied to a variety of problems, ranging from edge detection to high level object recognition. Probabilistic relaxation is an iterative approach such that at each iteration step contextual information from further afield is utilised to refine the labelling result. One could say that at each iteration step information conveyed by coarser resolutions of the scene is incorporated in the labelling process. Seen that way, probabilistic relaxation seems to work in the opposite direction from standard multiresolution/multiscale techniques. Indeed, in such methods, the coarse resolution information is extracted first and it is then used to guide the solution at the finer resolution. This direction of information flow seems to be more intuitive as it fits with the way we perceive scenes: we see something with our peripheral, low resolution vision and subsequently we foveate on it; we see something from a distance, with low resolution, and subsequently we approach and see it better, in more detail. In this paper we use such a causal approach to the problem of multiscale information incorporation. In particular, we develop a general framework for a probabilistic relaxation scheme that works *across* the various resolution levels and allows the incorporation of measurements pertaining at all levels. The scheme is used for

*ERASMUS student on visit from RWTH Aachen, Germany the segmentation of coloured texture images and involves the introduction of an appropriate dictionary.

Before we proceed further, we wish to clarify the use of the terms "multiresolution" and "multiscale" in the context of this work. The resolution of an image corresponds to the size of a pixel in physical units. As objects have fixed physical sizes, when a pixel represents a larger physical length, fewer pixels are needed to represent the same object. This is the intuitive basis of multiresolution pyramid creation: successive smoothings and subsamplings. However, if one does not perform the subsamplings, but only the successive smoothings, one may say that the approach is multiscale as the filters used to create the coarser versions of the image become larger and larger in pixel size, creating images that are more and more blurred but have the same number of pixels. That is the approach we use here and the reason we do it is to preserve the correspondence between the pixels across the different levels. Thus, we use the terms multiscale and multiresolution interchangeably, although the terms multilevel or even multiblur would be more appropriate.

In Section 2 we shall present a brief review of multiresolution relaxation schemes developed and used by various researchers. In Section 3 we shall develop our methodology. In Section 4, the process of extraction of multiscale features used in the segmentation of coloured textures shall be described, followed by a presentation of some results of our segmentation approach. We shall conclude in Section 5.

2 Literature Survey

Probabilistic relaxation has been used by a few researchers in multi-resolution representations of data. A major reason for adopting such an approach has been to reduce the computational load through a faster convergence. Glazer [2] applied "multi-level" relaxation to the Gauss-Seidel method of iterative solution of a finite-difference approximation for optic flow analysis. He used a multi-resolution pyramid representation of the input image. The iterative relaxation was effectively applied by formulating problems in low level vision as partial differential or finite difference equations. In a similar vain, Terzopoulos [11] developed multi-resolution iterative algorithms for other image analysis tasks, such as shape from shading, which are put forward mathematically as variational principles or as partial differential equations. Zhang et al. [13] applied a relaxation technique for edge detection on a multi-resolution hierarchical representation of the input image.

Hancock et al. [3] presented a probabilistic relaxation method for more consistent edge labelling but using a multi-resolution, hierarchical representation of the edge map. In this scheme, the probabilistic relaxation approach is applied to each level of resolution as if it were the single level approach, with the only difference that the "prior probabilities" of the labels at each level are adjusted to be consistent with the priors of the previous level as extracted from the dictionaries. Our scheme differs in the sense that it applies across different levels and it uses the same dictionary of allowable label configurations at all levels. Furthermore, we have a tower of images smoothed at different scales rather than a pyramid where the physical dimensions of the image are reduced.

Such multiscale approaches for texture analysis are few and far between. Unser and Eden [12] extracted texture energy measures form the image and smoothed the output of the extraction filter bank using Gaussian smoothing at different scales. The features in these multiscale planes are reduced, by diagonalising scatter matrices evaluated at two different spatial resolutions, and thresholded to yield texture segmentation. Matalas et al. [5], used a B-spline transform in order to obtain images at several smoothing levels to calculate vector dispersion and gradient orientation at different scales. A small disparity function was then applied to segment textures.

3 Multiscale Probabilistic Relaxation

The problem of multiscale probabilistic labelling of the input image using a set of blurred versions of the image can be stated as follows. Let l indicate the levels of coarsening with $l = 1, ..., L$, representing the levels from full resolution to the coarsest level. Let $i, i = 1, ..., N$ be a pixel and $\mathbf{x}_i^l$, the associated measurement vector for that pixel at resolution level l. We define a label set $\Omega, \Omega = \{\omega_1, \omega_2, ..., \omega_m\}$, which contains all possible labels of the image for m possible categories. Thus, each pixel i has label θ_i that can take on values from Ω.

We wish to choose for pixel i the most probable label θ_i given all the available information. In other words we wish to set:

$$\theta_i = arg\{\max_k P(\theta_i = \omega_k \mid \mathbf{x}_j^l, \forall j, \forall l)\} \tag{1}$$

For simplicity and clarity of exposition we shall restrict ourselves in considering only two successive levels of resolution l and $l+1$. Then, using Bayes's Rule:

$$P(\theta_i = \omega_k \mid \mathbf{x}_j^l, \mathbf{x}_j^{l+1}, \forall j) = \frac{P(\theta_i = \omega_k, \mathbf{x}_j^l, \mathbf{x}_j^{l+1}, \forall j)}{P(\mathbf{x}_j^l, \mathbf{x}_j^{l+1}, \forall j)} \tag{2}$$

We can expand the terms in the numerator and the denominator by applying the theorem of total probability:

$$P(\theta_i = \omega_k \mid \mathbf{x}_j^l, \mathbf{x}_j^{l+1}, \forall j) =$$
$$\frac{\sum_{\omega_{\theta_1}} .. \sum_{\omega_{\theta_{i-1}}} \sum_{\omega_{\theta_{i+1}}} .. \sum_{\omega_{\theta_N}} P(\theta_1 = \omega_{\theta_1}, .., \theta_i = \omega_k, .., \theta_N = \omega_{\theta_N}, \mathbf{x}_j^l, \mathbf{x}_j^{l+1}, \forall j)}{\sum_{\omega_{\theta_1}} \sum_{\omega_{\theta_2}} .. \sum_{\omega_{\theta_i}} .. \sum_{\omega_{\theta_N}} P(\theta_1 = \omega_{\theta_1}, .., \theta_i = \omega_{\theta_i}, .., \theta_N = \omega_{\theta_N}, \mathbf{x}_j^l, \mathbf{x}_j^{l+1}, \forall j)} \tag{3}$$

The joint probability that appears in Equation 3 can be factorised as follows:

$$\begin{aligned} &P(\theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}, \mathbf{x}_j^l, \mathbf{x}_j^{l+1}, \forall j) = \\ &\quad P(\mathbf{x}_1^{l+1}, \mathbf{x}_2^{l+1}, ..., \mathbf{x}_N^{l+1} \mid \theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}, \mathbf{x}_1^l, \mathbf{x}_2^l, ..., \mathbf{x}_N^l) \times \\ &\quad P(\theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}, \mathbf{x}_1^l, \mathbf{x}_2^l, ..., \mathbf{x}_N^l) \end{aligned} \tag{4}$$

As we try to emulate here perceptual segmentation, we can imagine that due to causality, measurements obtained at level $l+1$ (the coarser level) can not possibly depend on measurements obtained at level l. Thus, the first factor on the right hand side of Equation 4 can be simplified as follows:

$$\begin{aligned} &P(\mathbf{x}_1^{l+1}, \mathbf{x}_2^{l+1}, ..., \mathbf{x}_N^{l+1} \mid \theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}, \mathbf{x}_1^l, \mathbf{x}_2^l, ..., \mathbf{x}_N^l) = \\ &\quad P(\mathbf{x}_1^{l+1}, \mathbf{x}_2^{l+1}, ..., \mathbf{x}_N^{l+1} \mid \theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}) \end{aligned} \tag{5}$$

We also expect that the measurement concerning a certain pixel depends on the identity of that pixel alone and on nothing else. Therefore, we can further write:

$$\begin{aligned} &P(\mathbf{x}_1^{l+1}, \mathbf{x}_2^{l+1}, ..., \mathbf{x}_N^{l+1} \mid \theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}) = \\ &\quad \prod_j P(\mathbf{x}_j^{l+1} \mid \theta_j = \omega_{\theta_j}) = \prod_j \frac{P(\theta_j = \omega_{\theta_j} \mid \mathbf{x}_j^{l+1})}{\hat{p}(\theta_j = \omega_{\theta_j})} \hat{p}(\mathbf{x}_j^{l+1}) \end{aligned} \tag{6}$$

where $\hat{p}(\mathbf{x}_j^{l+1})$ is the prior probability of measurements $\mathbf{x}_j^{l+1}$ to arise, and $\hat{p}(\theta_j = \omega_{\theta_j})$ is the prior probability of label ω_{θ_j}.

Now consider the second factor on the right hand side of Equation 4:

$$\begin{aligned} &P(\theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}, \mathbf{x}_1^l, \mathbf{x}_2^l, ..., \mathbf{x}_N^l) = \\ &\quad P(\mathbf{x}_1^l \mid \theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}, \mathbf{x}_2^l, ..., \mathbf{x}_N^l) \times \\ &\quad P(\theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}, \mathbf{x}_2^l, ..., \mathbf{x}_N^l) \end{aligned} \tag{7}$$

We can further expand the second term on the right hand side of Equation 7 to write:

$$\begin{aligned} &P(\theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}, \mathbf{x}_1^l, \mathbf{x}_2^l, ..., \mathbf{x}_N^l) = \\ &\quad P(\mathbf{x}_1^l \mid \theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}, \mathbf{x}_2^l, ..., \mathbf{x}_N^l) \times \\ &\quad P(\mathbf{x}_2^l \mid \theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}, \mathbf{x}_3^l, ..., \mathbf{x}_N^l) \times .... \times \\ &\quad P(\mathbf{x}_N^l \mid \theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}) \times \\ &\quad P(\theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}) \end{aligned} \tag{8}$$

For the same reasons explained earlier, we expect that the measurement obtained for a particular object depends on the identity of the object itself and on nothing else. Thus, all factors on the right hand side of Equation 8, except the last one, can be simplified to express dependence only on the identity of the object they refer to. The last factor is the joint probability of a certain label assignment to arise. So we have:

$$\begin{aligned} &P(\theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}, \mathbf{x}_1^l, \mathbf{x}_2^l, ..., \mathbf{x}_N^l) = \\ &\quad \prod_j P(\mathbf{x}_j^l \mid \theta_j = \omega_{\theta_j}) \times P(\theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}) \end{aligned} \tag{9}$$

Now by substituting from Equations 6 and 9 in Equation 4:

$$\begin{aligned}
&P(\theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N}, \mathbf{x}_j^l, \mathbf{x}_j^{l+1}, \forall j) = \\
&\quad \prod_j \frac{1}{\hat{p}(\theta_j = \omega_{\theta_j})} \; P(\theta_j = \omega_{\theta_j} \mid \mathbf{x}_j^{l+1}) \; \hat{p}(\mathbf{x}_j^{l+1}) \; P(\mathbf{x}_j^l \mid \theta_j = \omega_{\theta_j}) \; \times \\
&\quad P(\theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N})
\end{aligned} \tag{10}$$

Then, upon substitution in Equation 3:

$$\begin{aligned}
&P(\theta_i = \omega_k \mid \mathbf{x}_j^l, \mathbf{x}_j^{l+1}, \forall j) = \\
&\frac{P(\mathbf{x}_i^l \mid \theta_i = \omega_k) \; P(\theta_i = \omega_k \mid \mathbf{x}_i^{l+1}) \; \hat{p}(\mathbf{x}_i^{l+1}) \; \prod_j \hat{p}(\mathbf{x}_j^{l+1}) \; Q(\theta_i = \omega_k)}{\sum_{\omega_{\theta_i}} P(\mathbf{x}_i^l \mid \theta_i = \omega_{\theta_i}) \; P(\theta_i = \omega_{\theta_i} \mid \mathbf{x}_i^{l+1}) \; \hat{p}(\mathbf{x}_i^{l+1}) \; \prod_j \hat{p}(\mathbf{x}_j^{l+1}) \; Q(\theta_i = \omega_{\theta_i})}
\end{aligned} \tag{11}$$

where

$$\begin{aligned}
&Q(\theta_i = \omega_{\theta_i}) = \\
&\quad \frac{1}{\hat{p}(\theta_i = \omega_{\theta_i})} \sum_{\omega_{\theta_1}} \cdots \sum_{\omega_{\theta_{i-1}}} \sum_{\omega_{\theta_{i+1}}} \cdots \sum_{\omega_{\theta_N}} \prod_j \frac{P(\mathbf{x}_j^l \mid \theta_j = \omega_{\theta_j}) P(\theta_j = \omega_{\theta_j} \mid \mathbf{x}_j^{l+1})}{\hat{p}(\theta_j = \omega_{\theta_j})} \; \times \\
&\quad P(\theta_1 = \omega_{\theta_1}, \theta_2 = \omega_{\theta_2}, ..., \theta_N = \omega_{\theta_N})
\end{aligned} \tag{12}$$

In the above expression $\hat{p}(\mathbf{x}_i^{l+1})$ is independent of the summation indices and cancels in the numerator and denominator. Therefore, the above expression further simplifies and gives the following formula for the labelling of a pixel from one level to the next:

$$\begin{aligned}
&P(\theta_i = \omega_k \mid \mathbf{x}_j^l, \mathbf{x}_j^{l+1}, \forall j) = \\
&\quad \frac{P(\mathbf{x}_i^l \mid \theta_i = \omega_k) \; P(\theta_i = \omega_k \mid \mathbf{x}_i^{l+1}) \; Q(\theta_i = \omega_k)}{\sum_{\omega_{\theta_i}} P(\mathbf{x}_i^l \mid \theta_i = \omega_{\theta_i}) \; P(\theta_i = \omega_{\theta_i} \mid \mathbf{x}_i^{l+1}) \; Q(\theta_i = \omega_{\theta_i})}
\end{aligned} \tag{13}$$

At the finest resolution equations 1 and 13 give the final segmentation result. The latter expression states that the probability of a certain pixel to have a certain label is equal to the product of the probability of the values of the measurements concerning the pixel to arise given the identity of the pixel, times the probability of the pixel to have this identity given the measurements concerning it at coarser resolutions, times a factor that expresses the contextual information. The denominator in the above formula is simply a normalising constant. The contextual information factor not only contains the product of measurement probabilities to arise given pixel identities, times probabilities of pixel identities given coarser measurements summed over all other pixels, but also the probability of a particular combination of pixel identities to arise. In the next section the meaning of all these factors will become clearer.

4 Application

Formula 12 is a sum over all possible label configurations of all pixels other than the pixel under consideration. Thus it is a formula of exponential complexity as it includes a very large number of terms. It can be simplified if we introduce a dictionary of permissible label configurations. As our scheme is aimed at performing region segmentation, we introduce a dictionary of 3×3 image patches where the labels of up to two different regions are allowed to co-exist. In Figure 1 we illustrate the entries of this dictionary. Letters A and B stand for two different region labels. Therefore, each entry of the dictionary depicted represents as many entries as there are pair combinations of the available region labels. For example, for 3 possible labels the dictionary will contain $13 \times 6 = 78$ possible allowable configurations.

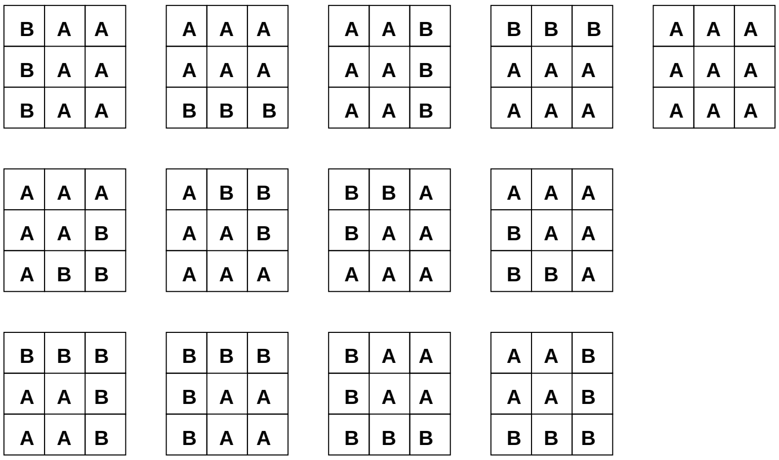

Figure 1: Entries in pattern dictionary

Further, the expression $P(\theta_j = \omega_{\theta_j} \mid \mathbf{x}_j^{l+1})$ is the probability of label ω_{θ_j} given to pixel j at the previous level of resolution. We can write this as $P^{l+1}(\theta_j = \omega_{\theta_j})$. In the absence of any prior knowledge, the prior probabilities of the various labels can be considered as constant and omitted from 12 because they would appear in both the numerator and the denominator of formula 13 and thus ultimately cancel. In view of the above, formula 12 simplifies to:

$$Q^l(\theta_i = \omega_{\theta_i}) = \sum_{\substack{\text{all patterns}\\ \text{in the}\\ \text{dictionary}}} \sum_{\substack{\text{all possible}\\ \text{combinations}\\ \text{of two labels}\\ \text{where centre}\\ \text{pixel has}\\ \text{label } \omega_{\theta_i}}} \prod_{\substack{3 \times 3 \text{ pixel}\\ \text{neighbourhood}\\ \text{entry in}\\ \text{dictionary}}} P(\mathbf{x}_j^l \mid \theta_j = \omega_{\theta_j}) P^{l+1}(\theta_j = \omega_{\theta_j}) \qquad (14)$$

$P(\mathbf{x}_j^l \mid \theta_j = \omega_{\theta_j})$ expresses the probability of a certain measurement to arise in each class. For example, in the problem we are applying this methodology, the measurements we use at each level of resolution are the values of the Luv colours of a pixel and the probability of each triplet of Luv values to arise is computed from the colour histogram of each class. The colour histogram of a class at resolution level l is created by reading the values of the pixel associated with this class with confidence higher than 80% at level $l+1$. The various levels of resolution are created by smoothing the image using the masks proposed by Zhang and Wandell [14]. These masks imitate the smoothing performed by the human vision system when it sees colour textures from various distances. Thus, the coarsest level we start from mimmicks the way the image would look when viewed from several meters distance, with all colour textural detail lost by blurring. This over-blurred image is clustered with the help of a K-means algorithm and the process of probabilistic relaxation is performed as we proceed towards the finer levels. We perform the clustering and the histogramming in the *Luv* colour space since the Euclidean differences between pixels in this space reflect their perceptual similarity. The initial probabilities are set using the clusters in the most blurred image. They are computed as a function of the squared distance of a pixel from the mean of each and every cluster [8].

Figure 2 displays the results of this approach on some collages of real colour textures of ceramic tiles. For comparison, we also show the result of segmentation achieved by Matas and Kittler [6] whose method of colour segmentation groups colour pixels by taking into account simultaneously both their feature space similarity and spatial coherence. Our segmentation accuracy calculated as a percentage of correctly segmented pixels, varied between 98.3 to 99.998% and these figures were consistently better than Matas and Kittler's approach.

5 Conclusions

We presented in this paper a general framework for applying the method of probabilistic relaxation across different resolution levels: the method allows the incorporation of information and measurements pertaining to separate resolution levels, and proceeds from the coarsest level to the finest level. As such, it allows the flow of information in a more natural way than the classical probabilistic relaxation approach where in successive iteration steps the information from coarser and coarser scales is incorporated. We applied our formalism to the problem of multiresolution segmentation of coloured texture images where a dictionary of permissible region label configurations is introduced; the method then proceeds in a bootstrap way by calculating the measurements that characterise a certain class from the pixels that have been assigned with high confidence to this class at the previous resolution level. The method was demonstrated with the help of some composite colour texture images where the measurements at each resolution level are the Luv colours that make up each texture at that level.

The basic assumptions the theory relies on is the independence of measurements concerning the individual pixels. This assumption is reasonably correct when the measurements are the colour/grey values of the pixels. It is clearly incorrect when the measurements refer to other pixel attributes like for example texture. However, it is an important convenience assumption, necessary to allow the development of the work. Often in practice, probabilistic label assignments are iterated to compensate for the error introduced by this assumption. For example, in our case, the relaxation labelling could be iterated a few times at each resolution level before the solution obtained was propagated to the next finer resolution level. A stopping criterion then would be necessary, which could for example be: the process stops when the number of pixels changing their probabilities are a small fraction of total image pixels.

References

[1] W.J. Christmas, J. Kittler, and M. Petrou. Structural matching in computer vision using probabilistic relaxation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 17(8):749–764, August 1995.

[2] F. Glazer. Multilevel relaxation in low-level computer vision. In A. Rosenfeld, editor, *Multiresolution Image Processing and Analysis*, pages 312–330. Springer, 1984.

[3] E. R. Hancock, M. Haindl, and J. Kittler. Multiresolution edge labelling using hierarchical relaxation. In *Proceedings of International Conference on Pattern Recognition*, pages 140–144, 1992.

[4] E.R. Hancock and J. Kittler. Edge labelling using dictionary-based relaxation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12:165–181, 1990.

[5] I. Matalas, S. Roberts, and H. Hatzakis. A set of multiresolution texture features suitable for unsupervised image segmentation. In *Proceedings of Signal Processing VIII, Theories and Applications*, volume III, pages 1495–1498, 1996.

[6] J. Matas and J. Kittler. Spatial and feature based clustering: Applications in image analysis. In *CAIP95*, pages 162–173, 1995.

[7] S. Peleg. A new probabilistic relaxation scheme. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2(4):362–369, July 1980.

[8] M. Petrou, M. Mirmehdi, and M. Coors. Perceptual segmentation of colour textures. Technical report, University of Surrey VSSP-TR-3/97, 1997.

[9] A. Rosenfeld, R. Hummel, and S. Zucker. Scene labeling by relaxation operations. *IEEE Trans. Systems, Man, and Cybernetics*, 6:420–433, June 1976.

[10] L.G. Shapiro and R.M. Haralick. Structural description and inexact matching. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 3:504–519, September 1981.

[11] D. Terzopoulos. Image analysis using multigrid relaxation methods. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8(2):129–139, 1986.

[12] M. Unser and M. Eden. Multiresolution feature extraction and selection for texture segmentation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11(7):717–728, July 1989.

[13] D. Zhang, J. Liu, and F. Wan. Multiresolution relaxation: Experiments and evaluations. In *Proceedings of International Conference on Pattern Recognition*, pages 712–714, 1988.

[14] X. Zhang and B.A. Wandell. A spatial extension of CIELAB for digital color image reproduction. In *Society for Information Display Symposium, San Diego*, 1996. WWW address: ftp://white.stanford.edu/scielab/spie97.ps.gz.

Caption for Figure 2 on next page: (left column) Original images, (middle column) Segmentation by applying Matas & Kittler's approach, and (right-column) Segmentation by multilevel probabilistic relaxation. Original images are in colour where different colours simply indicate different labels.

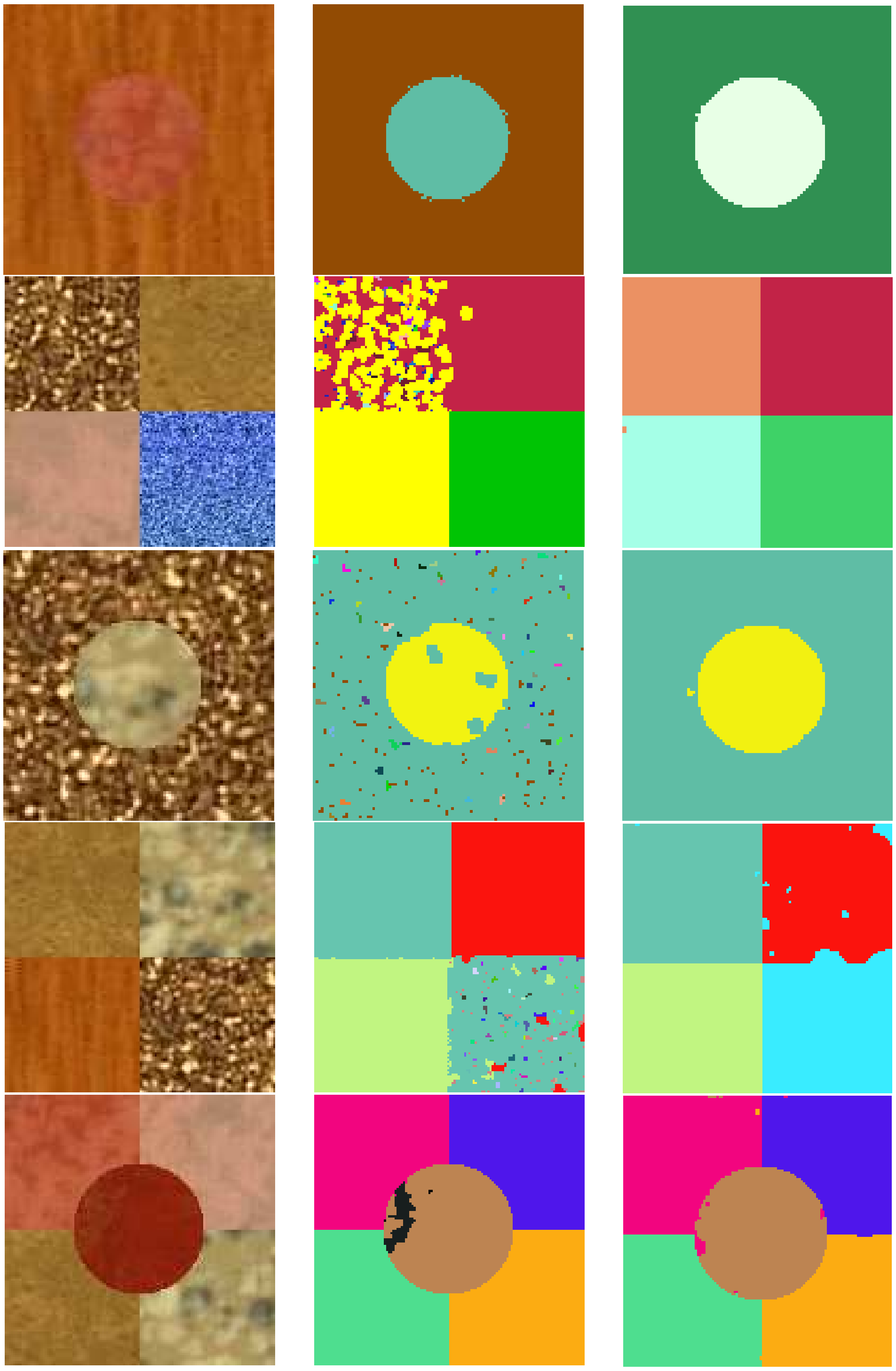

Figure 2: (for caption please see previous page).