Unsupervised image texture segmentation with optimized filters

Vidar Alvestad and Trygve Randen
Høgskolen i Stavanger
P.O. Box 2557 Ullandhaug, 4004 Stavanger, Norway
Email: vidar-a@hsr.no, tranden@hsr.no

ABSTRACT

This paper propose a new unsupervised image texture segmentation algorithm based on the design of general optimal FIR filters.

A pyramidal decomposition divide the image into equal dyadic regions/cells. A new similarity function is developed to compare adjacent cells. Similar cells are equally labeled through a labeling process yielding a number of distinct areas. Optimal filters are then computed with respect to each area. The final segmentation result is obtained with a local energy function followed by the c-means algorithm. Recommandations for further improvement of the algorithm are supplied.

Experiments shows that general optimal FIR filters are capable to produce satisfactory segmentation results in an unsupervised texture segmentation system.

1. INTRODUCTION

A vast amount of information about our environment is obtained through our visual system. A lot of this information concerns textures (patterns). This is why humans have developed a tremendous skill with which to identify regions with uniform textures in a given image, known as texture segmentation. This capability is desired to automate in many computer vision systems, such as: Inspection [1], medical image analysis [2, 3], document processing [4, 5], remote sensing [6], and compression [7].

In designing such systems, the intuitive objective is to achieve a system which do not need any operator guidance. This requires that we automatically segment the textured image in absence of any a-priori information, known as unsupervised texture segmentation.

Hundreds of segmentation techniques are present in the literature. The optimal filter design approach concerns the determination of a filter that provides the "largest" discrimination between two textures. One optimal filter technique involves determining optimal Gabor filter parameters with respect to some criterion function [8]. Teuner et.al. [9] have recently introduced an approach for the design of optimal Gabor filters based on the analysis of a spectral Gabor pyramid. However, recent research have shown promising results in designing general optimal filters not restricted to the Gabor class [10]. (Both a review of segmentation techniques and a reimplementation of the approach by Teuner et.al. [9] is presented in [11].)

2. THE PYRAMIDAL DECOMPOSITION

The idea in this new approach to unsupervised texture segmentation is to perform a coarse segmentation at first, using a similar pyramidal decomposition of the image as in the paper by Teuner [9]. This decomposition divides the image into equal dyadic regions/cells. Instead of analyzing each cell for spectral features, we will look at the surrounding cells for similarities. If neighboring cells are similar, they are merged. Therefore, functions that measure the similarity between adjacent cells are developed and tested. Optimal filters are then designed with respect to the merged areas. After the filtering process a local energy function is computed similar to the one in [12] which yields a set of feature images that will be clustered/segmented by the c-means algorithm.

2.1. Design of general optimal FIR filters

The objective of an optimal filter for a two texture discrimination is to achieve a good feature separation. In the paper by Randen and Husøy [12] they propose to use the Fisher criterion [13] to achieve this.

Several other object functions are proposed [14, 15, 16] , but as stated by Randen et.al. [12], their major drawback is that they do not incorporate the feature variances.

An interesting finding is done by Randen et.al. [12], the closed form solution of several optimization problems all entail the same eigenproblem:

$$(\mathbf{R}_{xx_i}^{-1}\mathbf{R}_{xxj})\mathbf{h} = \lambda\mathbf{h} \tag{1}$$

where $\mathbf{R}_{xxi}$ and $\mathbf{R}_{xxj}$ are constructed from the autocorrelation matrices corresponding to textures $x_i(m,n)$ and $x_j(m,n)$, respectively. In Equation

(1) λ represents the eigenvalues of $(\mathbf{R}_{xx_i}^{-1}\mathbf{R}_{xxj})$ and $\mathbf{h}$ the corresponding eigenvectors. Every vector $\mathbf{h}$ is a candidate filter which is composed of the lexicographically ordered rows of the two dimensional filter $h(m,n)$. The optimal filter is then selected as the eigenvector $\mathbf{h}$ yielding maximum object function.

2.2. Pyramidal decomposition

If we denote the image $x(m,n)$ of size N, the pyramidal decomposition splits the image into P cells of equal dyadic size:

$$x(m,n) = \bigcup_{i=1}^{P} x_i(m,n) \tag{2}$$

The number of cells at each level l is $P = 2^l$, and each cell size N_c is $N_c = \frac{N}{2^l}$.

2.3. Similarity function

If we could use an optimal filter framework in achieving a function that could decide whether or not two cells have similar textures, this certainly would be pleasant. We could then use equal formulas at both the coarse segmentation and later for refinement by optimal filter design.

Since the textures are modeled by their autocorrelation matrices it is natural to assume that similar textures have almost equal autocorrelation matrices and that different textures have not.

If we let $\mathbf{R}_{xxi} = \mathbf{R}_{xxj}$, all the eigenvalues in Equation (1) becomes equal. On the other hand, if we assume that $\mathbf{R}_{xxi}$ and $\mathbf{R}_{xxj}$ contains autocorrelation estimates of two different textures, the eigenvalues are expected to be more spread.

Therefore, in comparing two cells it is assumed to be sufficient to compare the eigenvalues spread in Equation (1). A large spread indicates that two cells contain two different textures, and little spread indicates two cells with similar textures.

Hence, the similarity function $d_{i,j}$ between texture cell $x_i(m,n)$ and $x_j(m,n)$ can be expressed as a function of the eigenvalue set $\{\lambda\}$, $d_{i,j} = f(\{\lambda\})$.

2.4. Properties of the eigenvalues

From linear algebra, the characteristic equation from a matrix can both have real and complex roots, hence the eigenvalues obtained can both be real and complex in addition to the zero solution.

Since $\mathbf{R}_{xxi}$ is symmetric, $\mathbf{R}_{xx_j}^{-1}$ is symmetric, $(\mathbf{R}_{xxi}\mathbf{R}_{xx_j}^{-1})$ is symmetric and hence the eigenvalues of $(\mathbf{R}_{xxi}\mathbf{R}_{xx_j}^{-1})$ are real. By Section 3.1 of [17], it is shown that the eigenvalues of $(\mathbf{R}_{xxi}\mathbf{R}_{xx_j}^{-1})$ may be expressed by $\frac{\mathbf{h}^T\mathbf{R}_{xxi}\mathbf{h}}{\mathbf{h}^T\mathbf{R}_{xxj}\mathbf{h}}$ of the smoothed *squared* filter outputs from filter $\mathbf{h}$. Due to the squaring operator, this is inherently a non-negative number. Hence, the eigenvalues are all non-negative.

As a consequence of this $(\mathbf{R}_{xx_i}^{-1}\mathbf{R}_{xxj})$ is positive semidefinite so that its eigenvalues are all nonzero and real.

2.5. Eigenvalue spread

Two calculations of the eigenvalue spread are performed.

If the maximum and minimum eigenvalues are λ_{max} and λ_{min} the eigenvalue contrast function ξ_{CT} is defined as,

$$d_{i,j} = \xi_{CT} = \frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}}.$$

However, measuring the spread as a function of only two of a set of eigenvalues is expected to reveal only partly the behavior of the eigenvalue spread.

A function that takes the whole eigenvalue set into consideration is the ratio of the geometric mean to the arithmetic mean ξ_{AG}:

$$d_{i,j} = \xi_{AG} = \frac{\frac{1}{N}\sum_{i=1}^{N}\lambda_i}{(\prod_{i=1}^{N}\lambda_i)^{\frac{1}{N}}}.$$

This function is equal to unity if all eigenvalues are equal, (indicating cells containing equal textures) and approaches zero as the eigenvalue spread increases (indicating different texture cells).

A threshold value is introduced to uniqly decide whether two cells contains similar textures or not:

$$\text{if} \begin{cases} d_{i,j} < \xi_{AG,t} & \text{, two cells are different} \\ d_{i,j} > \xi_{AG,t} & \text{, two cells are similar} \end{cases} \tag{3}$$

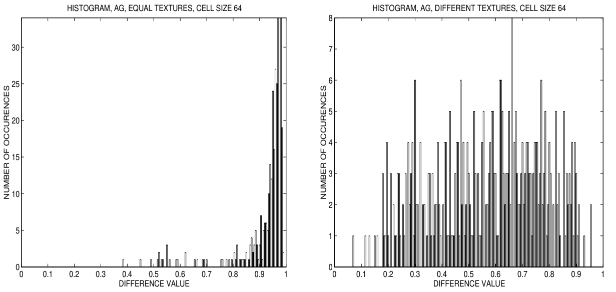
In Equation (3) only the threshold value corresponding to the AG function $\xi_{AG,t}$ is shown. In the next subsection we will develop a way for finding the optimal threshold.

2.6. Optimal threshold values

The problem of deciding whether or not two cells are similar can be formulated as a two class pattern recognition problem. Based on the value of $d_{i,j}$ we shall decide whether or not the cells $x_i(m,n)$ and $x_j(m,n)$ are similar.

Let w_1 be the class when we have two cells containing similar textures, and w_2 the class when they are different. Since we do not know anything about the probability of occurrence of the two classes, a reasonable assumption is that the a priori probabilities are equal, that is $P(w_1) = P(w_2)$.

The behavior of the similarity value, $d_{i,j}$, is dependent on to which class it is referred to. This is expressed in terms of the class conditional probability density functions (pdf), $p(d_{i,j}|w_1)$ and $p(d_{i,j}|w_2)$.



(a) (b)

Figure 1. (a) and (b) Histograms, AG measure, cell size 64. (a) Cells containing similar textures. (b) Cells containing different textures.

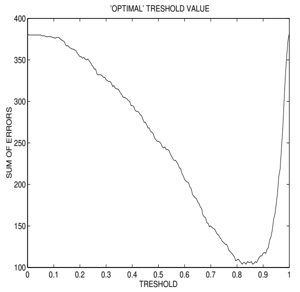


Figure 2. Classification error, AG measure, cell size 64, optimal threshold, $\xi_{AG,t} = 0.845$.

The strategy of finding an optimal threshold value is approached by a consideration of classification error, which is defined as $(d_{i,j} = \xi_{AG,t})$:

$$P(e) = P(w_1)\int_0^{\xi_{AG,t}} p(d_{i,j}|w_1)\mathbf{d}d_{i,j} + P(w_2)\int_{\xi_{AG,t}}^1 p(d_{i,j}|w_2)\mathbf{d}d_{i,j}$$

The optimal threshold value is then chosen as the one minimizing the error function $P(e)$. Since $P(w_1) = P(w_2)$, this can be formulated as:

$$\min_{\xi_{AG,t}} \{P(e)\} = \min_{\xi_{AG,t}} \left\{ \int_0^{\xi_{AG,t}} p(d_{i,j}|w_1)\mathbf{d}d_{i,j} + \int_{\xi_{AG,t}}^1 p(d_{i,j}|w_2)\mathbf{d}d_{i,j} \right\}$$

We must estimate the distributions $p(d_{i,j}|w_1)$ and $p(d_{i,j}|w_2)$ to obtain an optimal threshold value. The functional form of the distributions are assumed unknown, and a non-parametric approach which use histograms is chosen.

The data set is taken from the Brodatz album [18], a total of 21 different images were selected.

N_c	$\xi_{AG,t}$	$\xi_{CT,t}$	$P_{AG}(e)$	$P_{CT}(e)$
128	0.935	0.665	33	46
64	0.845	0.770	104	130
32	0.825	0.920	134	167

Table 1. Experimental test.

The samples which corresponds to class w_2 (different textures) are obtained by computing the similarity function between each of the different textures. We do not have such a number of similar textures. And because the data set should incorporate the same set of textures, the samples which correspond to class w_1 (similar textures) are obtained by sliding a window an equal number of times in each of the textures and computing the similarity functions. Both functions are computed, and for cell size $N_c = 128, 64$ and 32, i.e. pyramid level 1, 2 and 3.

The estimated pdf's are obtained by computing the histograms with the similarity measure $d_{i,j}$ discretizised to 200 levels. The histograms corresponding to the cell size $N_c = 64$ are plotted in Figure 1 (a) and (b). The shapes are as expected in the previous discussion.

The discrete solution to the optimal threshold value is found by plotting the error function $P(e)$ with respect to the threshold value and selecting the minimum. The classification error corresponding to the AG function is plotted in Figure 2.

The threshold values together with corresponding minimal errors, $P_{AG}(e)$ and $P_{CT}(e)$, are outlined in Table (1). Three observations are noted: 1) The similarity functions have different optimal threshold values at the three pyramid levels. 2) The classification error increase as the cell size decrease (as we could intuitively expect). 3)The classification error $P(e)$ is significant smaller for the AG function.

Consequently, the AG function is selected as the measure of similarity, and the threshold values $\xi_{AG,t}$ are selected to make the decisions at the different pyramid levels.

2.7. Label connected cells

A cell $x_i(m,n)$ has four horizontal or vertical neighbors and four diagonal neighbors. The horizontal and vertical neighbors are called the 4-neighbors, and the set of all neighbors are called the 8-neighbors of image cell $x_i(m,n)$. Thus, a complete comparison of the surrounding cells involves computing eight values.

We could use the set of 4-neighbors, but 8-neighbors have more connections to the surrounding cells, which hopefully has the effect of more homogeneous areas.

At a certain initial chosen level, the image is scanned cell by cell from left to right and from top to bottom. Let the neighboring cells be as depicted

$x_o(m,n)$	$x_p(m,n)$	$x_q(m,n)$
$x_r(m,n)$	$x_i(m,n)$	

Figure 3. Neighbors of a cell.

in Figure 3. The nature of the scanning sequence ensures that when we come to $x_i(m,n)$, the cells $x_o(m,n), x_p(m,n), x_q(m,n), x_r(m,n)$ have already been labeled. If all neighbors have similarity values below $\xi_{AG,t}$, the cell $x_i(m,n)$ is not equal to its surrounding cells, and is therefore assigned a new label. If two or more neighbors have similarity measurements above $\xi_{AG,t}$, thus indicating equal cells, one of the labels is assigned to $x_i(m,n)$ and a note of the appropriate equivalences is made. If only one neighbor indicate equality, this label is assigned to the cell.

After a complete scan of the image, the equivalent label pairs are sorted into equivalence classes and unique labels are assigned to each class. A second scan of the image is then performed, replacing each label by the label assigned to its equivalence class.

2.8. Merge cells

The net result of the previous step is a labeled map. Cells which have equal labels are merged, yielding a number of distinct regions in the image. If a region contains just one cell, we assume that this cell is a result of a transition between neighboring regions, and this cell is omitted.

2.9. Optimal filters

The optimal filters are designed with respect to each of the different labeled areas, as shown in Section 2.1. Thus, for each labeled area, we compute optimal filters with respect to all other labeled areas.

2.10. Segmentation

The final segmentation result is achieved by using the set of optimal filters yielding a set of feature images after the local energy function [12]. Each of the feature images are used to form feature vectors where each feature image correspond to one dimension in the feature space. Feature vectors extracted from similar textures should lie close to each other in feature space forming a cluster. In present implementation the c-means algorithm is used to find the clusters and thus the segmented result.(A complete outline of the experimental system is given in [11].)

3. DISCUSSION/RESULTS

The nature of the labeling process ensures a more robust algorithm. The number of cells is assumed to be considerably larger than the number of texture regions in the image. Thus, the a priori probability of similar cell *neighbors* is higher than the a priori probability of different cell *neighbors*. The nature of the labeling sequence take this into consideration and give preference to homogeneous areas: All visited neighbors must be considered as different if a new label is assigned, but only one label is sufficient to incorporate the cell as similar to its surrounding cells. The situation becomes more complex if two or more visited cells have different labels and are detected as similar. A texture cell can not be similar to two texture cells that are different(!). This implementation consider this as a wrong decision in a homogeneous area, and all the labels are considered equivalent.

A function that performs a good and robust measure of the similarity is crucial in the system design, since the system design rely on this performance.

Further, the approach depends essentially on the following initial chosen parameters:

The threshold value ξ_t. This value is important in achieving the areas which the optimal filters are designed with respect to. If a wrong decision is made, this directly influence the optimal filters performance. Optimal ξ_t are computed a priori with the thechnique outlined in Section 2.6.

The size of the general optimal FIR filter. As shown by Randen et.al. [17], a fairly small size (8×8) filter is able to produce appreciable feature separation. This size is expected to be appropriate for performance evaluation, and further optimization is therefore not done.

The pyramid level l, to which we analyze the image. The cell size N_c should ideally be small to superimpose the smallest texture area in the image. On the other hand, it must be large to be able to incorporate the texture features.

3.1. Examples

3.1.1. Example Nat2

The Nat2 texture mosaic contains two textures (D55 and D56) of rather similar appearance. The texture image is shown as a faint background in the label map as depicted in Figure 4.

This example shows that several of the cells in the texture boundaries are detected as different to their surroundings. Cells labeled 3 through 7 are all single and therefore rejected. The segmented result at this pyramid level ($l = 3$) is shown in Figure 5.

3.1.2. Example Nat3

This texture consist of a rectangular area (D17) in the center of a background texture region (D6).

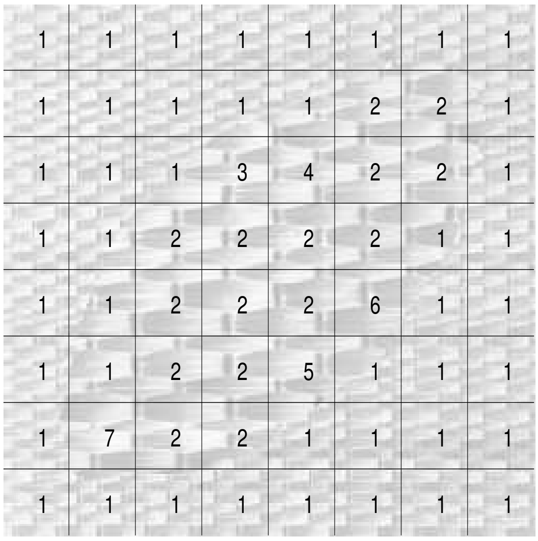


Figure 4. Label map of image Nat2 (level 3).

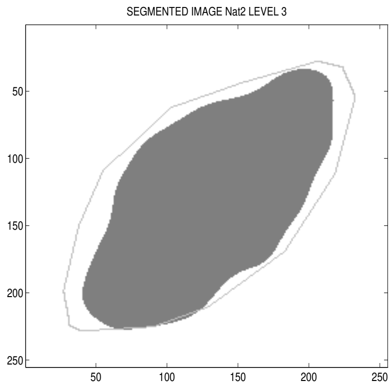


Figure 5. Segmented result of image Nat2 (level 3).

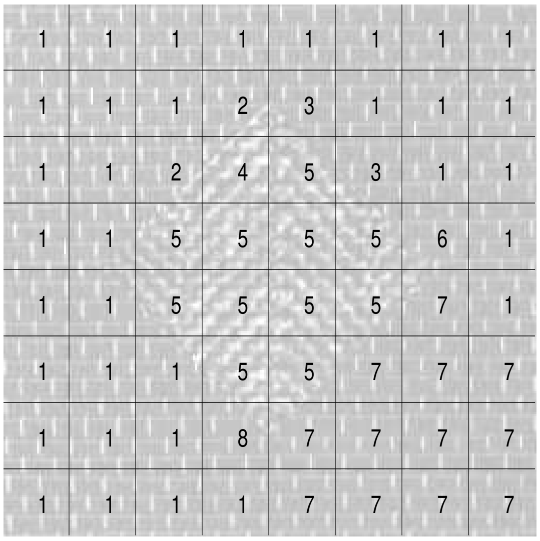
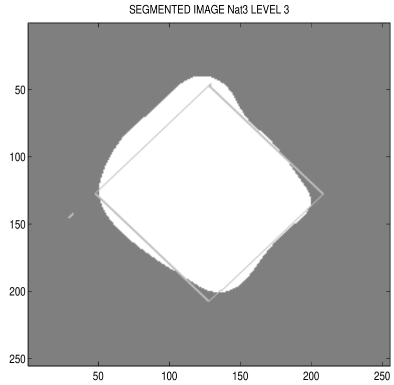


Figure 6. Label map of image Nat3 (level 3).



(b)

Figure 7. Segmented result of image Nat3 (level 3). Labeled region 1 and 7 have been merged before computing optimal filters.

The label map of Figure 6, shows that the background is wrongly detected as two different areas, labeled area 1 and 7. If we compute the similarity function on the basis of these areas, they are detected as similar since the new computation incorporate a larger data set and also contain little of the center texture relative to the areas. This yields the appreciable result depicted in Figure 7. The small regions are considered as transitions between texture regions and consequently omitted.

This clearly indicate that the similarity computation should incorporate a second merging step. Future development of the algorithm should therefore consider:

- After computing the similarity between small cells at a high pyramid level, the algorithm could be rerun computing the similarity between cell clusters.
- The possibility to incorporate the relativeness between the similarity values in the labeling process.
- Any cell size can be considered, and cell sizes are not restricted to be equal.

4. CONCLUSION

The new idea of this unsupervised texture segmentation approach is the computation of the similarity measure to roughly estimate the texture regions which is used in the optimal filter design. The computation of the eigenvalue spread is shown to be a suitable tool to measure the similarity between texture regions. The examples show that the general optimal FIR filters are capable to produce satisfactory segmentation results.

REFERENCES

[1] M. Tuceryan and A. K. Jain, "Texture analysis," in *Handbook of pattern recognition and*

computer vision (C. H. Chen, L. F. Pau, and P. S. P. Wang, eds.), ch. 2, pp. 235–276, Singapore: World Scientific, 1993.

[2] S. Kjøde, "Optimal filters for detection of stellate- and circumscribed lesions in mammograms," Master's thesis, Høgskolen i Stavanger, Norway, 1995. To be published.

[3] S. O. Gabrielsen, "Automated classification of microcalcifications in mammograms using a multichannel filtering approach," Master's thesis, Høgskolen i Stavanger, Norway, 1995. To be published.

[4] A. K. Jain and S. Bhattacharjee, "Text segmentation using Gabor filters for automatic document processing," *Machine Vision and Applications*, vol. 5, no. 3, pp. 169–184, 1992.

[5] T. Randen and J. H. Husøy, "Segmentation of text/image documents using texture approaches," in *Proc. NOBIM-konferansen-94*, (Asker, Norway), pp. 60–67, June 1994.

[6] E. Rignot, R. Chellappa, and P. Dubois, "Unsupervised segmentation of polarimetric SAR data using the covariance matrix," *IEEE*, vol. 30, pp. 697–705, July 1992.

[7] M. R. El-Sakka and M. S. Kamel, "A segmentation criterion for digital image compression," *IEEE*, pp. 2551–2554, 1995.

[8] D. F. Dunn, W. E. Higgins, and J. Wakeley, "Texture segmentation using 2-D Gabor elementary functions," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, pp. 130–149, Feb. 1994.

[9] A. Teuner, O. Pichler, and B. J. Hosticka, "Unsupervised texture segmentation of images using tuned matched Gabor filters," *IEEE Trans. Image Processing*, vol. 4, pp. 863–870, July 1995.

[10] T. Randen and J. H. Husøy, "Optimal texture filtering," in *Int. Conf. on Image Proc.*, (Washington D.C.), Oct. 1995.

[11] V. Alvestad, "Unsupervised image texture segmentation with optimized filters," Master's thesis, Høgskolen i Stavanger, Norway, 1995. To be published.

[12] T. Randen and J. H. Husøy, "Design of optimal filters for image texture feature separation," in *Proc. NORSIG*, pp. 221–226, Sept. 1995.

[13] K. Fukunaga, *Statistical Pattern Recognition*. San Diego: Academic Press, 2 ed., 1990.

[14] H. Singh and A. Mahalanobis, "Correlation filters for texture recognition and applications to terrain-delimination in wide-area surveillance," in *Proc. Int. Conf. Acoust. Speech, Signal Proc.*, pp. V153–V156, IEEE, 1994.

[15] A. Mahalanobis and H. Singh, "Application of correlation filters for texture recognition," *Applied Optics*, vol. 33, no. 11, pp. 2173–2179, 1994.

[16] M. Unser, "Local linear transforms for texture measurements," *Signal Processing*, vol. 11, no. 1, pp. 61 – 79, 1986.

[17] T. Randen and J. H. Husøy, "Optimal filters for two-class texture feature extraction," *Pattern Recognition*, 1995. submitted.

[18] P. Brodatz, *Textures: A photographic album for artists and designers*. NY: Dover, 1966.