HIERARCHICAL TEXTURE SEGMENTATION USING DICTIONARIES

P. Bajcsy and N. Ahuja
University of Illinois at Urbana-Champaign
Beckman Institute
405 N. Mathews, Urbana, IL 61801

Abstract

We present a new hierarchical texture segmentation method that partitions an image into textured regions. A textured region is viewed as a set of uniformly distributed primitives. A primitive is a region with constant gray values. Gray values within a primitive can be corrupted by noise. Any noisy primitive contains gray values from a δ-wide interval (δ-homogeneous primitive). The noisy primitive is described by the mean of interior gray values. A textured region with noise is characterized by a set of gray value means (texture dictionary) derived from noisy primitives. Every pixel (sample point) and its neighborhood give rise to an estimate of texture dictionary. Components of the estimated dictionary at a pixel characterize noisy primitives of a textured region grown from the pixel. Co-occurrence of noisy primitives from this grown region are calculated. Final segmentation is obtained by grouping pixels with identical dictionaries and co-occurrences created at each pixel. Homogeneity degree δ of noisy primitives provides a framework for multiscale analysis. Computational efficiency and robustness of the proposed method are related. Experiments are reported for synthetic and real textures from Brodatz album and real scenes.

1 Introduction

The goal of any image segmentation is to partition an image (grid of samples) into connected subsets of samples, denoted as regions, each having a uniform texture. Textures have no universal model. A variety of texture models have been derived from: (1) perceptual studies [9, 8, 6] (mimicking humans), (2) specific two-dimensional tasks [13], such as, automated surface inspection (textile, paint), medical image processing (semi-automated search for tissues, tumors), (3) texture gradient analyses [3, 2] (projective distortion problem) and (4) texture imitation [10, 5] (realism in computer graphics). Texture segmentation methods have been based on two types of models: statistical [7, 5] and structural methods [14, 12], which can be alternately viewed as pixel-based and region-based models [11]. The motivation for the work described in this paper is to develop a robust and computationally efficient hierarchical texture segmentation method.

Following are some salient characteristics of the work presented in this paper. (1) It demonstrates satisfactory segmentation performance of the statistical-structural approach to texture modeling ([14, 7, 1]) (2) It uses second order (co-occurrence) statistics of regions [7] that are themselves derived by grouping pixels and are therefore unrestricted in the size and shape they assume; in contrast co-occurrence has mostly been considered for pixel properties in the past work. (3) It provides a mechanism for tradeoff between computational requirements and robustness of the segmentation obtained. (4) Segmentation is hierarchical and autonomous; the hierarchical segmentation derived is indexed by a scale parameter.

Our work here takes a structural-statistical approach. First in Sec. 2, an idealized textured region C_k is modelled as a set of uniformly distributed, arbitrary shaped primitives. A primitive is a region with constant gray values[1]. Gray values within a primitive can be corrupted by noise. Any noisy primitive contains gray values from a δ-wide interval (δ-homogeneous primitive). The noisy primitive is characterized by the mean of interior gray values. A textured region C_k with noise is characterized by a set of gray value means (texture dictionary $\vec{v}_k$), which is the set of the means of the noisy primitives. Observable (realistic) textures are derived from the ideal textures by creating perturbations in the (1) primitive model, (2) primitive distribution and (3) amount of noise in primitives. Other issues, such as, variable size of texture dictionary and partially different dictionaries of two textures (dictionary overlaps), are considered.

Given an image and the texture model, a new tex-

[1] Gray value is later replaced with the term attribute.

ture segmentation method is described in Sec. 3. First, the maximum possible size of texture dictionary is estimated. This size of dictionaries guarantees that a global behavior of textures is captured. Second, texture dictionaries are estimated at every pixel (sample point) from the neighborhood of a pixel. Components of the estimated dictionary at a pixel characterize noisy primitives of a textured region grown from the pixel. This grown region is a pixel driven, global, spatially irregular estimation of a textured region. Third, co-occurrence of primitives from this grown textured region are calculated. Lastly, a final segmentation is obtained by grouping pixels with identical dictionaries and co-occurrences of corresponding primitives. Robustness against variations of the texture model observed in realistic textures is analyzed. Computational efficiency and robustness of the proposed method are related.

In Sec. 4, multiple texture segmentations are performed for various degrees of homogeneity δ (called scale parameter) within a primitive. A hierarchy of detected regions is identified by allowing regions only to grow for increasing values of δ.

Sec. 5 shows performance evaluation on synthetic images (robustness and efficiency) and Brodatz and real scene textures. Concluding remarks are stated in Sec. 6.

2 Modelled and Observable Textures

A certain degree of mathematical formalism is needed before the segmentation method is proposed in Sec 3.

2.1 Image and segmentation

Input data, a two-dimensional image, is represented by a function $x_i \longrightarrow f(x_i)$, where x_i is a sample point $x_i = (x_1, x_2)$ in the domain of function f and $f(x_i)$ is the function value denoted as attribute at each location pointed to by sample point x_i. The goal of texture image segmentation is to partition the sample points x_i in an image into sets of connected sample points, denoted as regions, such that each region consists of uniformly distributed primitives.

2.2 Texture model

Let us consider an image segmented into piecewise constant regions S_j (identical attributes inside S_j). Let us assume that every textured region C_k consists of a set of primitives equal to the subset of region S_j, $sub\{S_j\}$ (see Figure 1 top).

Each textured region C_k can be characterized by a set of attributes $f(x_i)$ present in $sub\{S_j\}$, e.g., three gray values $f_1 = v1$, $f_2 = v2$ and $f_3 = v3$ of the region C_1 shown in Figure 1 (top). This set of attributes for a textured region C_k is denoted by a texture dictionary

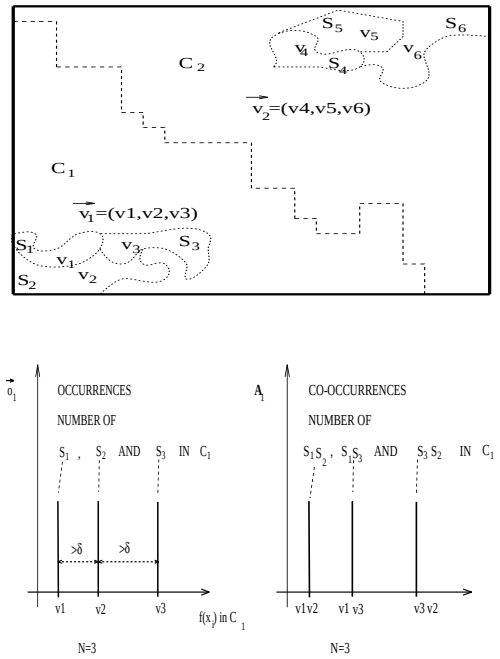

Figure 1: Textured region and its occurrence and co-occurrence of primitives.

$\vec{v}_k = (v1, v2, ...)$ containing N_k entries (a size of the dictionary is N_k). Two textured regions have different dictionaries.

Each element from a texture dictionary $\vec{v}_k$ characterizes multiple primitives $S_j \in C_k$. The number of primitives $S_j \in C_k$ characterized by each dictionary component creates a component of an occurrence vector $\vec{o}_k$. The number of neighboring primitive pairs in C_k gives rise to a component of a co-occurrence matrix $\mathbf{A_k} = \{\mathbf{a_{i,j}}\}$, e.g., a pair (S_1, S_2) occurs in the region C_1 $a_{1,2}$-times. Figure 1 (bottom) shows occurrence and co-occurrence values of region C_1, where components on the horizontal axis are labeled with the dictionary values ($v1, v2, ..$ or $v1v2, ...$). The diagonal elements of $\mathbf{A_k}$ are equal to zero. A textured region consists of uniformly distributed primitives therefore (1) components of occurrence vectors and (2) off-diagonal components of co-occurrence matrix have similar values (ideally values are identical).

A texture primitive with noise is modelled as a δ-homogeneous region and obtained by performing δ-homogeneous image segmentation. δ-homogeneous image segmentation partitions the image into a set of regions S_j with δ homogeneity, where j is the index of a region. The homogeneity δ of one region S_j^δ [2] is

[2] S_j^δ - denotes a region S_j with δ homogeneity. δ can take all possible values from 0 up to the maximum attribute value.

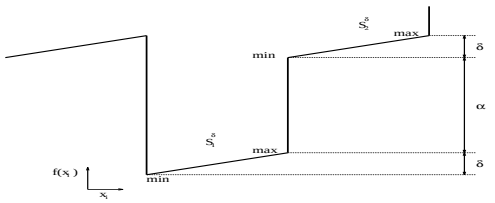

Figure 2: Homogeneity δ and contrast α definitions.

defined as the maximum distance between all pairs of attributes (e.g., gray values) at the sample points from the region S_j^δ; $\| f(x_i \in S_j^\delta) - f(x_k \in S_j^\delta) \| \leq \delta$ (see Figure 2). The contrast α of two neighboring regions S_1^δ and S_2^δ is defined as the minimum distance between all pairs of attributes at one sample point from each of the two regions; $\alpha = \alpha_{1,2} = \min\{ \| f(x_i \in S_1^\delta) - f(x_k \in S_2^\delta) \| \}$ (see Figure 2).
The noisy primitive is described by the mean of its interior attributes (first-order statistics). A textured region C_k with noise is characterized by a set of attribute means (texture dictionary $\vec{v}_k$), which is the set of the means of the noisy primitives. A distribution of primitives can be observed from the co-occurrences of noisy primitives (second-order statistics) characterized by texture dictionary.

2.3 Observable textures

Observable (realistic) textures possess properties either directly captured by the texture model or achieved by several variations of the texture model[3]. Variations of the following ideal characteristics are considered (illustrations of variations are in Figure 3):
(1) Model for a primitive.
(2) Distribution of primitives within a textured region (Figure 3; periodic - left column; random uniform - top middle and right, bottom middle and right; random nonuniform - middle right).
(3) Random noise inside of a primitive (Figure 3, top right, bottom right).
Two more issues related to dictionaries in observable textures are shown in Figure 3; a variable size of dictionaries (bottom row) and partially different dictionaries of neighboring textured regions - dictionary overlaps (middle left, bottom row).

The developed texture segmentation method based on the texture model provides correct segmentation (shown in the center of Figure 3) for all images in Figure 3. Other issues related to illumination conditions (e.g., shadows, highlights, blur) are not addressed.

[3] Observable textures could be also defined as transformations of an ideal periodic texture according to Zucker [14].

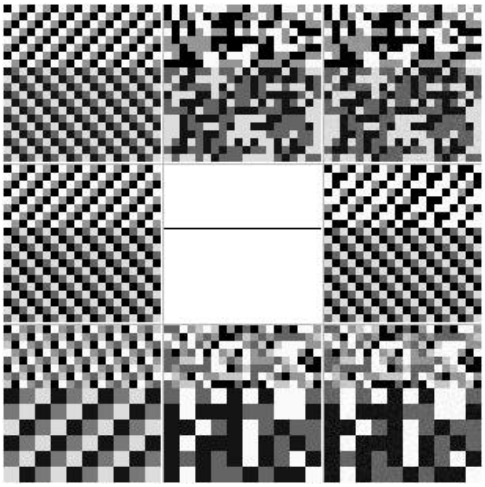

Figure 3: Variations of a texture model.

3 Texture Segmentation

First, the proposed method is described. It is followed by a discussion of (A) computational efficiency and (B) robustness against (1) noise in primitives, which effects dictionary detection, (2) nonuniform distribution of primitives tightly related to dictionary overlaps.

3.1 The texture segmentation method

Assumptions: Let us consider that a given image contains δ-homogeneous regions S_j^δ (texture primitives) and all neighboring pairs of regions S_j^δ have contrast $\alpha > \delta$. Let us assume that a subset $sub_1\{S_j^\delta\}$ of all regions S_j^δ creates an unknown connected textured region $C_1^\delta = sub_1\{S_j^\delta\}$. The textured region C_1^δ is characterized by (1) N_1 attribute means of texture primitives (components of a dictionary $\vec{v}_1$), which are mutually more than δ apart and (2) identical values of occurrence components $\vec{o}_1$ and co-occurrence off-diagonal components $\mathbf{A_1}$ due to a uniform distribution of primitives.
Derived textured region at each sample point: Given a fixed size of dictionary N, a dictionary $\vec{v}_{x_i}$ can be built at each sample point x_i by searching in a neighborhood of x_i[4]. The first component of the dictionary is $v1 = f(x_i)$. The next value of the dictionary is any attribute, which is more than δ apart from all dictionary elements found before; e.g., $v2 = f(x_k)$ if $\| f(x_i) - f(x_k) \| > \delta$.

From the dictionary $\vec{v}_{x_i}$ found at each sample point x_i, a 2δ-homogeneous textured region[5] $C_{x_i}^{2\delta}$ is created

[4] x_i became an index of the dictionary $\vec{v}_{x_i}$ since every sample point creates its own dictionary.

[5] Notation: Derived textured region $C_{x_i}^{2\delta}$ is found at each

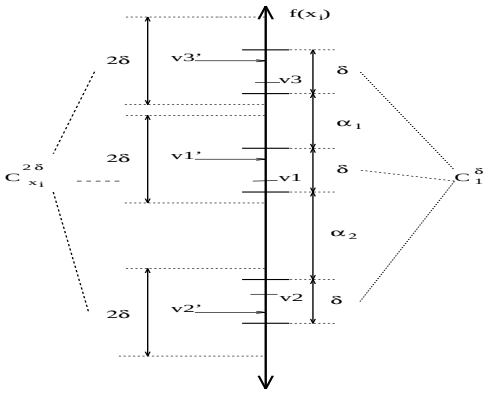

Figure 4: Dictionaries $\vec{v}_1 = (v1, v2, v3)$ and $\vec{v}_{x_i} = (v1', v2', v3')$ for regions C_1^δ and $C_{x_i}^{2\delta}$ in a case of $N = N_1 = 3$.

by grouping together all neighboring sample points x_k[6] satisfying the inequality $cyc\{\| \ \vec{v}_{x_i} - \vec{v}_{x_k} \ \|\} \leq \delta$, where $cyc\{.\}$ represents subtraction of any cyclical variation of dictionary components. A co-occurrence matrix $\mathbf{A_{x_i}}$ of primitives S_j^δ within $C_{x_i}^{2\delta}$ is calculated.

Analysis based on size of texture dictionary: Three types of textured regions $C_{x_i}^{2\delta}$ exist with respect to a priori unknown textured region C_1^δ depending on a selected dictionary size N.

If the assumed size N of a dictionary is identical to the actual dictionary size N_1, $N = N_1$, then a priori unknown region C_1^δ is identical with all regions $C_{x_i}^{2\delta}$ found, where x_i belongs to the interior of C_1^δ. In this case, dictionaries for a calculated region $C_{x_i}^{2\delta}$ and for a priori unknown region C_1^δ satisfy the inequality $cyc\{\| \ \vec{v}_1 - \vec{v}_{x_i} \ \|\} \leq \delta$. The dictionary components are shown in Figure 4.

If $N > N_1$ then all N_1 components of the dictionary $\vec{v}_1$ (corresponding to a priori unknown region C_1^δ) are not more than δ apart from N_1 out of N components of all dictionaries $\vec{v}_{x_i \in C_1^\delta}$ found. The N_1 components from the dictionaries $\vec{v}_{x_i \in C_1^\delta}$ are identified by selecting dictionary components having identical values of their off-diagonal co-occurrence components from $\mathbf{A_{x_i}}$ (Figure 5). Adjacent samples x_i and x_l belong to the same textured region if (1) subsets of dictionary components satisfy the following inequality $cyc\{\| \ sub\{\vec{v}_{x_i}\} - sub\{\vec{v}_{x_l}\} \ \|\} \leq \delta$ and (2) a co-occurrence component of $(v1_{x_i}, v1_{x_l})$ has identical value with all values of off-diagonal co-occurrence components from $\mathbf{A_{x_i}}$ and $\mathbf{A_{x_l}}$ over these subsets (see Figure 5).

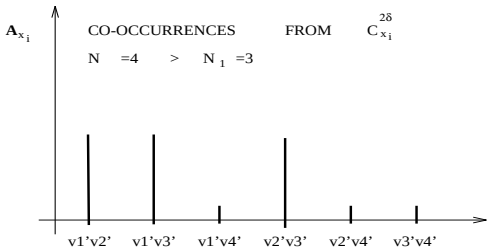

Figure 5: Distribution of off-diagonal co-occurrence matrix components.
The co-occurrence matrix $\mathbf{A_{x_i}}$ is created from $C_{x_i}^{2\delta}$ using a dictionary $\vec{v}_{x_i} = (v1', v2', v3', v4')$ of the size $N = 4 > N_1 = 3$.

If $N < N_1$ then the dictionary $\vec{v}_{x_i}$ as well as the co-occurrence matrix $\mathbf{A_{x_i}}$ cannot estimate global properties of a textured region.

Size of dictionary: Following from the previous analysis, the goal is to select a dictionary size N greater or equal to the unknown true dictionary size N_1. Any unknown size N_1 of dictionary $\vec{v}_1$ is bounded by the minimum and maximum values, $N_{min} \leq N_1 \leq N_{max}$. The minimum size of the dictionary is two, $N_{min} = 2$, because in order to create a textured region there must be at least two primitives with distinct characteristic attributes. The maximum size of the true dictionary N_{max} is derived considering two neighboring textured regions with non-overlapping dictionaries, which span the whole attribute range $R_{total} = N_1 * (\delta_1 + \alpha_1) + N_2 * (\delta_2 + \alpha_2)$. Then the maximum size of a dictionary is found by setting $N_1 = N_{min} = 2$ and $N_2 = N_{max}$; $N_{max} = \frac{R_{total} - 2*(\delta_1 + \alpha_1)}{(\delta_2 + \alpha_2)}$. The dependency of N_{max} on δ is plotted in Figure 6 for the range of gray scale images $R_{total} = 256$, $\delta_1 = \delta_2$ and $\alpha_1 = \alpha_2 = \delta_1 + 1$; $N_{max} = \frac{256}{2*\delta+1} - 2$. The size of a dictionary is selected to be $N = N_{max}$ if $N_{max} \geq N_{min}$ else $N = N_{min}$. Values of N can be only integer numbers in order to create dictionaries with discrete components.

Steps in the proposed texture segmentation method:
(1) For a fixed δ, estimate the maximum size of the true dictionary N_{max}.
(2) Find a dictionary $\vec{v}_{x_i}$ of the size $N = N_{max}$ at each sample point x_i consisting of attributes from a neighborhood of x_i, which are more than δ apart.
(2) Create textured regions $C_{x_i}^{2\delta}$ at each sample point

x_i (in the subscript) and by searching for samples around x_i with attributes $\pm\delta$ apart from dictionary components of $\vec{v}_{x_i}$, thus spanning 2δ wide attribute range (in superscript) by each dictionary component.

[6] x_k is neighboring to x_i if the distance of the mutual difference is equal to one; i.e., $\| x_k - x_i \| = 1$.

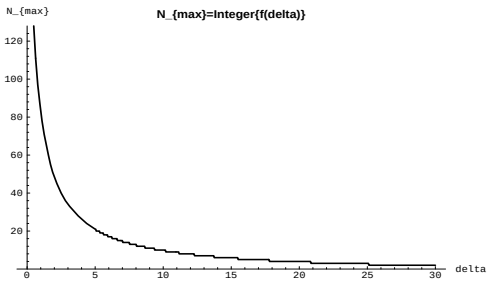

Figure 6: Maximum size of dictionary as a function of homogeneity δ.

x_i consisting of connected samples having attributes not more than δ apart from at least one component of the dictionary $\vec{v}_{x_i}$.
(3) Calculate co-occurrences $\mathbf{A_{x_i}}$ of primitives S_j^δ described by $\vec{v}_{x_i}$ within a textured region $C_{x_i}^{2\delta}$.
(4) Compare dictionaries $\vec{v}_{x_i}$, $\vec{v}_{x_l}$ and check uniformity of off-diagonal co-occurrence components $\mathbf{A_{x_i}}$, $\mathbf{A_{x_l}}$ for all adjacent samples x_i, x_l.
(5) Assign sample points x_i, x_l into final textured regions based on (4).

3.2 Computational efficiency

Computational efficiency is achieved by (1) lowering the dimensionality of computations to obtain dictionaries $\vec{v}_{x_i}$ together with derived regions $C_{x_i}^{2\delta}$ and (2) replacing co-occurrence matrices with co-occurrence vectors corresponding to regions $C_{x_i}^{2\delta}$.
Calculation of $\vec{v}_{x_i}$ and $C_{x_i}^{2\delta}$: Any one-dimensional cross section of the two-dimensional textured region C_1^δ can be used to estimate an unknown dictionary $\vec{v}_1$ by $\vec{v}_{1D,x_i}$ despite $\vec{v}_{x_i}$ such that the inequality $cyc\{\| \ \vec{v}_{1D,x_i} - \vec{v}_1 \ \|\} \leq \delta$ is always satisfied for $N = N_1$ and S_j^δ with $\delta < \alpha$. If the size of selected dictionary N is greater than an unknown size of texture dictionary N_1 then the inequality is always true in a form $cyc\{\| \ sub\{\vec{v}_{1D,x_i}\} - \vec{v}_1 \ \|\} \leq \delta$. Two dictionaries are used at each sample x_i to estimate an unknown dictionary $\vec{v}_1$ from two mutually perpendicular one-dimensional cross sections (rows and columns) starting from x_i.

If the primitives S_j^δ with $\delta < \alpha$ are uniformly distributed over a constituting textured region C_1^δ then any one-dimensional cross section (denoted as $C1D_{x_i}^{2\delta}$) of the two-dimensional textured region $C_{x_i}^{2\delta}$ will contain proportional numbers of occurrences and co-occurrences of primitives S_j^δ. Two one-dimensional regions $C1D_{x_i}^{2\delta}$ are created at each sample x_i to estimate co-occurrences along two mutually perpendicular directions. By lowering the dimensionality of calculations we decrease computational requirements.
Calculation of co-occurrence matrices $\mathbf{A_{x_i}}$: In order to merge two adjacent samples there is no need to calculate the whole co-occurrence matrix $\mathbf{A_{x_i}}$. It is sufficient to calculate at each sample x_i a co-occurrence vector $\vec{a}_{x_i}$, which contains all co-occurrences of $v1 = f(x_i)$ with remaining components of $\vec{v}_{x_i}$, $\vec{a}_{x_i} = (v1v2, v1v3, v1v4, ...)$. Co-occurrence vectors $\vec{a}_{x_i}$ contain the information about a distribution of the off-diagonal co-occurrence components and about a co-occurrence of two adjacent samples considered for a merger.

3.3 Noise robustness of dictionary detection

We assumed that the primitives S_j^δ had larger contrast α than homogeneity δ. In this case texture dictionaries are found correctly such that all dictionaries contain significant discriminable attributes of textured regions. There is no error in detection of textured regions due to incorrect components of dictionaries and calculated co-occurrences.

If $\delta \geq \alpha$ for a primitive S_j^δ due to a random zero-mean noise then there will be a finite error probability of an incorrect dictionary detection. In order to increase noise robustness of the dictionary detection, components of each dictionary $\vec{v}_{x_i}$ are updated after finding $C_{x_i}^{2\delta}$. The new dictionary components are means of attributes from repetitively occurring primitives $S_j^\delta \in C_{x_i}^{2\delta}$ described by the old dictionary components. If $\delta \geq \alpha$ then the error probability for the first dictionary component $v1_{x_i}$ can be expressed as $Pr(\| \ v1_{x_i \in S_j^\delta \subset C_1^\delta} - v1_1 \ \| \leq \frac{\delta}{2}) = \xi$, where $v1_1$ is the first dictionary component of $\vec{v}_1$ having a correct attribute mean based on primitives S_j^δ in C_1^δ and ξ is a confidence coefficient. The larger number of samples used for making the estimate $\vec{v}_{x_i}$ the more noise robust is the dictionary detection. As a consequence, there is always a tradeoff between efficiency and noise robustness since lowering dimensionality of calculations decreases the number of used samples for making an estimate $\vec{v}_{1D,x_i}$ as opposed to $\vec{v}_{x_i}$.

3.4 Nonuniform distribution and dictionary overlap

Properties of observable textures demonstrate a certain degree of deviation from the ideal uniform distribution of primitives. One would like to incorporate into the method robustness against nonuniform distribution of primitives (denoted as RN) and maximize RN.

Maximal RN means that a minimal number of components from cyclically matched dictionaries $\vec{v}_{x_i}$ and $\vec{v}_{x_l}$ is sufficient to merge samples x_i and x_l (assuming that corresponding co-occurrence components have identical value). This means that a small dictionary overlap will lead to erroneous merger since overlapped components of dictionaries will have identical corresponding co-occurrence components. Thus the robustness against dictionary overlaps RO is minimal. We can conclude that the following equation governs the two parameters RN and RO; $RO + RN = constant$.

4 Hierarchical Texture Segmentation

Texture segmentations for different values of a homogeneity parameter δ (scale parameter) create a multiscale space of segmentations. Increasing δ from small values to large values will increase probability of obtaining accurate primitives but decrease the discrimination power of the texture detection due to an increased probability of dictionary overlaps. Hierarchical segmentations are created by imposing a causality rule in a multiscale space. Hierarchy in multiscale space (causality rule) is formulated as follows: *Every region C_k^δ obtained at scale δ cannot split at $\delta + \triangle\delta$ into smaller subregions (bottom-up approach) and cannot merge at $\delta - \triangle\delta$ with other regions (top-down approach).*

The hierarchy in multiscale space provides better noise robustness for larger values of δ (larger δ means larger confidence interval thus smaller error probability). The hierarchy is guaranteed by modifying attribute values within created regions C_k^δ at each scale δ to the mean values of created regions based on their dictionaries $g(x_i, \delta) = \frac{1}{M_{j,k}} \sum_{l=1}^{M_{j,k}} f(x_l \in S_j^\delta \subset C_k^\delta)$, where $M_{j,k}$ is the number of samples in all primitives $S_j^\delta \subset C_k^\delta$ described by one dictionary component. 1

5 Performance Evaluation

Performance of the proposed texture segmentation method is judged based on (1) robustness (noise, nonuniform distribution of primitives, dictionary overlap), (2) computational requirements and (3) results from real textures (Brodatz textures, real scenes).

5.1 Robustness

Due to the relationship between robustness against dictionary overlaps and nonuniformity of primitives, two cases of robustness were separately tested: (1) noise robustness of dictionary detection and (2) robustness against nonuniform distribution and dictionary overlap. Experiments were conducted with synthetic images having all deviations from the texture model except the basic model for a primitive.

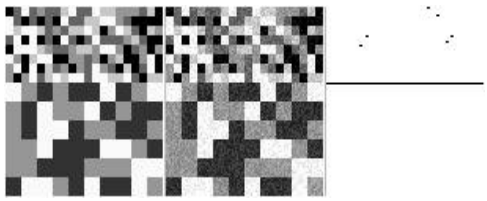

Figure 7: Segmentation of noisy primitives. Left - original, middle - noisy original with primitives having $\delta > \alpha = 0$, right - obtained segmentation for $N = 5$. Two textured regions C_1, C_2 with dictionaries $\{0, 100, 150, 200, 250\}$ and $\{50, 150, 250\}$ ($N_1 = 5$, $N_2 = 3$ and overlap= 2) corrupted by noise such that $\delta = 50$ and $\alpha = 1$.

Noise robustness of dictionary detection: Attributes $f(x_i)$ are corrupted by an additive noise $n(x_i)$ with a known symmetric probability distribution function D_n (zero mean, standard deviation σ), $p(x_i) = f(x_i) + n(x_i)$. By introducing a uniformly distributed noise we could control two crucial parameters δ and α of primitives S_j^δ as their mutual relationship $\delta = 0 < \alpha = 50$ continuously changed to $\delta = 50 > \alpha = 0$. There is no error in dictionary detection for $\delta < \alpha$. It was experimentally verified that there is only a small error for $\delta > \alpha = 0$ (see Figure 7 in a presence of dictionary overlaps) by observing the number of misclassified sample points (less than 2.5%).

Robustness against nonuniform distribution and dictionary overlap: Several experiments were conducted to test robustness against (1) noise in primitives, (2) dictionary overlap RO and (3) nonuniform distribution of primitives RN present in an image at once (shown in Figure 8). Relationship between (2) and (3) ($RO + RN = const$) required to choose an optimal solution (equal weights on RO and RN). To cope with dictionary overlaps (primitives from neighboring textured regions described by identical attribute means) along the border of two textured regions, one has to merge samples rather than primitives obtained from δ-homogeneous segmentation (see Figure 7 and 8).

5.2 Computational requirements

Computational requirements are proportional to the number of given sample points, the size of dictionary N and the number of sample points in $C_{x_i}^{2\delta}$ used for calculating co-occurrences. Computational time and memory are approximately linearly increasing with increasing size of dictionaries N. In average, one segmentation of a 2D gray scale image ($100x100$ samples) takes between $3 - 4s$ for $N = 3$ and $6 - 7s$

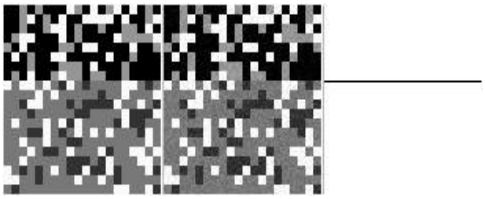

Figure 8: Segmentation of textures with nonuniform distribution of primitives.
Left - original: two textured regions C_1, C_2 with dictionaries $\{0, 150, 250\}$ and $\{50, 120, 250\}$ ($N_1 = N_2 = 3$ and overlap$= 1$), distribution of primitives with probabilities $\frac{15}{27}, \frac{2}{9}, \frac{2}{9}$ in C_1 and $\frac{2}{9}, \frac{15}{27}, \frac{2}{9}$ in C_2; middle - noisy original image with primitives having homogeneity $\delta = 30$; right - obtained segmentation for $N = 3$ from the noisy image.

for $N = 5$ on Sparc 20 workstations.

5.3 Experimental results

Due to the high computational requirements for the size of dictionary $N = N_{max}$ at small scales δ and some empirical observations, we ran the following segmentations with a fixed size $N < N_{max}$ of a dictionary over a range of scales δ. The final segmentation was selected manually from the hierarchy of segmentations.
Brodatz textures Experiments with textures from Brodatz album [4] were performed and shown in Figure 9 and 10. As long as there were perceptually different sets of gray scale values (attributes) the segmentation was performed correctly.
Real scene textures: The proposed texture segmentation method was applied to gray scale images of real scenes as a preprocessing step for an object recognition. Figure 11 contains a bear with fur and a checkerboard pattern vest segmented into three main regions (head, nose and vest), which turns out to be consistent with our perceptual expectation. A tree and grass in Figure 12 were expected to create two dominant textured regions (tree and its background holes, light and dark grass with ground). Grass shows approximately uniform distribution of its stems (varying in attributes due to lighting conditions) but the tree and its holes do not. Segmentation of the grass is successful but segmentation of a tree is only partially successful.

6 Summary and Conclusions

We have presented a new hierarchical texture segmentation method. The following characteristics of the work were shown:
(1) satisfactory segmentation performance of the statistical-structural approach to texture modeling (compare with [14]),

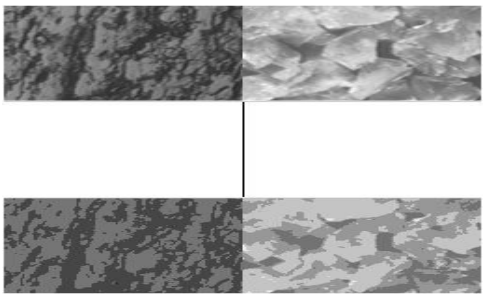

Figure 9: Brodatz texture I.
Top - original: Bark of tree (D12) and crushed rose quartz (D98); middle - segmentation using $N = 5$ at $\delta = 17$; right - reconstructed image $g(x_i, \delta = 17)$.

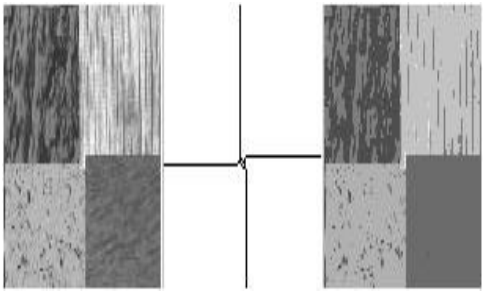

Figure 10: Brodatz texture II.
Left - original: Bark of tree (D12), wood grain (D68), pressed cork (D32) and water (D38); middle - segmentation using $N = 3$ at $\delta = 20$; right - reconstructed image $g(x_i, \delta = 20)$.

(2) direct use of co-occurrence probabilities derived from data driven, global, spatially irregular regions, (compare with other methods using co-occurrences introduced in [7]),
(3) tradeoff between computational requirements (proportional to dimensionality of computations) and noise robustness of the proposed method and
(4) unsupervised hierarchical segmentation providing results indexed by the scale parameter δ.
Results showing robustness against (1) noise in primitives and (2) nonuniform distribution of primitives were reported for synthetic data. Computational efficiency of the segmentation method was measured. Experiments with data from Brodatz album [4] and real scenes were conducted.

References

[1] N. Ahuja and B. J. Schachter. Image models. *Computing Surveys*, 13:373–397, Dec. 1981.

[2] R. Bajcsy and L. Lieberman. Texture gradient as a depth cue. *Computer Graphics and Image*

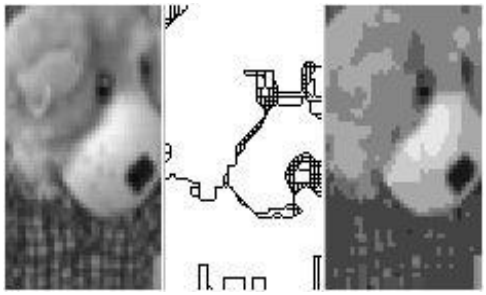

Figure 11: Bear from the image "Toys".
Left - original; middle - segmentation with $N = 3$ at $\delta = 12$; right - reconstructed image $g(x_i, \delta = 12)$.

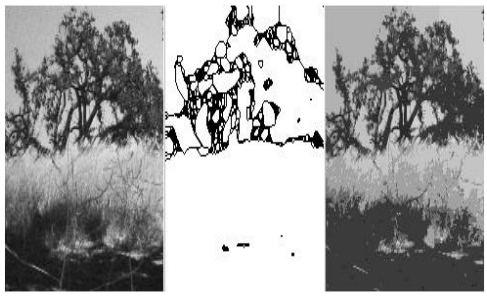

Figure 12: Tree and grass
Left - original; middle - segmentation with $N = 5$ at $\delta = 24$; right - reconstructed image $g(x_i, \delta = 24)$.

Processing, 5:52–67, 1976.

[3] D. Blostein and N. Ahuja. Shape from texture: Integrating texture-element extraction and surface estimation. *IEEE on PAMI*, 11(12):1233–1251, December 1989.

[4] P. Brodatz. *Textures: A photographic Album for Artists and Designers*. New York, NY, Dover, 1966.

[5] G. R. Cross and A. K. Jain. Markov random field texture models. *IEEE on PAMI*, 5(1):25–39, January 1983.

[6] O. D. Faugeras. Texture analysis and classification using a human visual model. In *Proc. of the 4th Int. Joint Conf. on Pattern Recognition*, pages 549–552, Kyoto, Japan, Nov. 1978.

[7] R. M. Haralick. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5):786–804, May 1979.

[8] B. Julesz. Experiments in the visual perception of texture. *Scientific American*, 232:34–43, April 1975.

[9] J. Malik and P. Perona. Preattentive texture discrimination with early vision mechanisms. *Journal of Optical Society of America*, 7(5):923–932, May 1990.

[10] A. P. Pentland. Fractal-based description of natural scenes. *IEEE on PAMI*, 6(6):661–674, November 1984.

[11] B. Schachter and N. Ahuja. Random pattern generation processes. *Computer Graphics and Image Processing*, 10:95–114, 1979.

[12] W. H. Tsai and K. S. Fu. Image segmentation and recognition by texture discrimination: A syntactic approach. In *Proc. of the 4th Int. Joint Conf. on Pattern Recognition*, pages 560–564, Kyoto, Japan, Nov. 1978.

[13] M. Tuceryan and A. K. Jain. *The Handbook of Pattern Recognition and Computer Vision, Eds. C. H. Chen and L. F. Pau and P. S. Wang*, chapter 2.1 Texture Analysis, pages 235–276. World Scientific Company, 1992.

[14] S. W. Zucker. Toward a model of texture. *Computer Graphics and Image Processing*, 5:190–202, 1976.