Unsupervised Texture Segmentation

Michal Haindl[1,2]

[1] Institute of Information Theory and Automation of the Czech Academy of Sciences,
[2] Center for Machine Perception, Czech Technical University,
Praha, Czech Republic
E-mail: haindl@utia.cas.cz

Abstract. A novel unsupervised multispectral texture segmentation algorithm is introduced. The textured image segmentation is based on a causal adaptive regression model prediction for detecting different types of texture segments which are present at the image. Texture segments are detected in four mutually perpendicular directions in the image lattice. Every monospectral component is checked separately and single monospectral results are combined together. The predictor in each direction uses identical contextual information from the pixel's neighbourhood and can be evaluated using a robust recursive algorithm. The method suggested can be successfully applied also to other unsupervised image segmentation applications, e.g. range image segmentation, edge detection, etc.

Keywords: Texture segmentation.

1 Introduction

Segmentation is a fundamental process affecting the overall performance of a machine vision system and texture is the important characteristics which exist in many natural images and so it plays an important role in both human and machine perception. Texture segmentation is a significant part of autonomous navigation systems, automatic inspection and quality control, virtual reality systems, etc. and as such it has been an active research area for past twenty years. Segmentation of a textured image should partion this image into meaningful patches representing different types of textures present at the visual scene. Usual approach is to assume knowledge of possible textures present in the scene, i.e. supervised segmentation. Our task is complicated by assuming none of such a prior knowledge. Besides its primary segmentation functionality an optimal segmentation algorithm should be also stable, accurate and numerically efficient.

There are many segmentation algorithms published in computer vision literature and a number of good survey articles [4],[6] is available. However the mutual comparison of segmentation algorithms is very difficult because of lack of sound experimental evaluation results together with available experimental data. The common approaches to solve the textured image segmentation problem are region growing based algorithms where single regions are formed by iteratively growing from seed regions, split-and-merge [5], clustering, and edge based techniques. Region growing and split-and-merge algorithms present the problem that they have to deal with different threshold values that are difficult to obtain and depend on an application. Clustering algorithms suffer usually less influence from this kind of problem, although other problems exist. Current segmentation algorithms most often miss small regions, detect false regions or suffer with over or under segmentation.

The present paper is organised as follows. In Section 2, a proposed method general concept under a Bayesian framework is introduced. Sections 3 and 4 complete the algorithm with an optimal model selection rule design and the voting criterion, respectively. The Section 5 deals with numerical realisation problems while Section 6 contains evaluation results obtained on a texture data test set.

2 Causal Monospectral Texture Model

We assume that spatial correlation characterising a region texture can be approximated by a spatial causal regression model. Unfortunately there is seldom preferred direction of dependence in textured images, i.e. texture data have usually noncausal dependence. Artificially imposing causal constraint neglects part of information present in data but leads to fast recursive algorithms of the Kalman-Bucy type what is the desired property in efficient segmentation applications. The regression model exploits high spatial correlation between neighbours of a modelled texture pixel. We neglect mutual spectral correlation and model each monospectral texture component independently. Although the recursive regression model can be easily generalized into the full 3D multispectral model, the chosen 2D simplification significantly reduces number of parameters to be estimated. We assume mono-spectral pixels to be modelled as:

$$Y_t = P^T Z_t + E_t \tag{1}$$

where $P^T = [a_1, \ldots, a_\beta]$ is the $1 \times \beta$ unknown parameter vector $\beta = card I_t$. We denote the $\beta \times 1$ data vector

$$Z_t = [Y_{t-i} : \forall i \in I_t]^T \tag{2}$$

with a multi-index $t = (m, n, d)$; Y_t is a predicted mono-spectral pixel value, m is the row number, n the column number, d $(d \geq 1)$ denotes the number of spectral bands, E_t is the white noise component. I_t is some neighbour index shift set such as the contextual neighbourhood is causal.

Note that although the model predicts a mono-spectral pixel component, the model can still use information from all other spectral bands in the case of a colour (multi-spectral) image $(d > 1)$. If the segmented textured image is monospectral then $d = 1$.

Let us choose a direction of movement on the image plane e.g. $t - 1 = (m, n-1, d), t-2 = (m, n-2, d), \ldots$ The white noise component E_t has zero mean and constant but unknown dispersion Ω. We assume that the probability density of E_t has a normal distribution independent of previous data and is the same for every time t. The task consists in finding the conditional prediction density $p(Y_t|Y^{(t-1)})$ given the known process history $Y^{(t-1)} = \{Y_{t-1}, Y_{t-2}, \ldots, Y_1, Z_t, Z_{t-1}, \ldots, Z_1\}$ and taking its conditional mean estimation $\tilde{Y}$ for the predicted data. If a prediction error is greater than an adaptive threshold the algorithm assumes crossing a border between two different texture regions. The conditional mean estimator was chosen as the predictor, because of its optimal properties ([1]):

$$\tilde{Y}_t = E[Y_t|Y^{(t-1)}] \tag{3}$$

Assuming normality of the white noise component E_t, conditional independence between pixels and an a priori probability density for the unknown model parameters chosen in the form (this normal form of a priori probability results in analytically manageable form of a posterior probability density)

$$p(P, \Omega^{-1}|Y^{(0)}) = (2\pi)^{-\frac{\gamma(0)}{2}} \; |\Omega|^{-\frac{\gamma(0)}{2}} \exp\left\{-\frac{1}{2}tr\{\Omega^{-1}\begin{pmatrix}-I\\P\end{pmatrix}^T V_0 \begin{pmatrix}-I\\P\end{pmatrix}\}\right\} , \tag{4}$$

where V_0 is a positive definite $(\beta+1)*(\beta+1)$ matrix and $\gamma(0) > d$, we have shown ([3]) that the conditional mean value is:

$$\tilde{Y}_t = \hat{P}^T_{t-1} Z_t \ . \tag{5}$$

The following notation is used in (4) and (5):

$$\hat{P}_{t-1} = V^{-1}_{zz(t-1)} V_{zy(t-1)} \ , \tag{6}$$

$$V_{t-1} = \tilde{V}_{t-1} + V_0 \ , \tag{7}$$

$$\tilde{V}_{t-1} = \begin{pmatrix} \tilde{V}_{yy(t-1)} & \tilde{V}^T_{zy(t-1)} \\ \tilde{V}_{zy(t-1)} & \tilde{V}_{zz(t-1)} \end{pmatrix} \ . \tag{8}$$

$$\tilde{V}_{xw(t-1)} = \alpha \tilde{V}_{xw(t-2)} + X_{t-1} W^T_{t-1}$$

It is easy to check (see [3]) also the validity of recursive (9). We assume slowly changing parameters, consequently these equations were modified using a constant exponential "forgetting factor" α to allow parameter adaptation.

$$\hat{P}_t = \hat{P}_{t-1} + (\alpha^2 + Z^T_t V^{-1}_{zz(t-1)} Z_t)^{-1} V^{-1}_{zz(t-1)} Z_t (Y_t - \hat{P}^T_{t-1} Z_t)^T \tag{9}$$

If the prediction error is larger than the adaptive threshold

$$|\tilde{Y}_t - Y_t| > \frac{1}{l} \sum_{i=1}^{l} |\tilde{Y}_{t-i} - Y_{t-i}| \tag{10}$$

then the pixel t is classified to a different texture class than its predecessor, i.e $\omega_t \neq \omega_{t-1}$ and the region border indicator ψ_t is set to 1. Otherwise both pixels share the same texture class $\omega_t = \omega_{t-1}$ and $\psi_t = 0$. The adaptive threshold is proportional to the local mean prediction error estimation.

3 Optimal Model Selection

The remaining problem is how to select an optimal contextual support set of the model (I_t). This can be done using the Bayesian theory. The following results can be either used to find an globally optimal model for the whole image before the segmentation step or to select a locally optimal model from a set of mutually competing models for each image lattice index during the segmentation process.

Let us assume several causal regression models (1) M_i with the number of unknown parameters· (β_i) and neighbour index shift sets $I_{t,i}$. The models may differ also in their forgetting factors α_i. According to the Bayesian theory, the optimal decision rule for minimizing the average probability of decision error chooses the maximum a posterior probability model, i.e. a model whose conditional probability given the past data is the highest one. The predictor used in the presented algorithm can be therefore completed for as in (11):

$$\tilde{Y}_t = \hat{P}^T_{j,t-1} Z_{j,t} \;\text{ if }\; p(M_j|Y^{(t-1)}) = \max_i\{p(M_i|Y^{(t-1)}) \tag{11}$$

where $Z_{j,t}$ is a data vector corresponding to the selected model. Following the Bayesian framework used in our paper and choosing uniform a priori model in the absence of contrary information, $p(M_i|Y^{(t-1)}) \sim p(Y^{(t-1)}|M_i)$, the simultaneous conditional probability density can be evaluated from

$$p(Y^{(t-1)}|M_i) = \int\int p(Y^{(t-1)}|P,\Omega^{-1}) p(P,\Omega^{-1}|M_i) dP d\Omega^{-1} \; . \tag{12}$$

Under the already assumed conditional pixel independence, the analytical solution has the form

$$p(M_i|Y^{(t-1)}) = k \, \frac{\Gamma\left(\frac{\gamma_i(t-1)-\beta_i+2}{2}\right)}{\Gamma\left(\frac{\gamma_i(0)-\beta_i+2}{2}\right)} |V_{i,zz(t-1)}|^{-\frac{1}{2}} \, \lambda_{i,t-1}^{-\frac{\gamma_i(t-1)-\beta_i+2}{2}} \; , \tag{13}$$

where k is a common constant. All statistics related to a model M_i (7)-(9), (13) are computed using the exponential forgetting constant α_i if the models differ also in their adaptation speed. The solution of (13) uses the following notations:

$$\gamma_i(t) = \alpha_i^2 \gamma_i(t-1) + 1 \; , \tag{14}$$

$$\lambda_{i,t-1} = V_{i,yy(t-1)} - V^T_{i,zy(t-1)} V^{-1}_{i,zz(t-1)} V_{i,zy(t-1)} \; . \tag{15}$$

The determinant $|V_{i,zz(t)}|$ as well as $\lambda_{i,t}$ can be evaluated recursively ([3]):

$$|V_{i,zz(t)}| = |V_{i,zz(t-1)}|\ \alpha_i^{2\beta_i}\ (1 + Z_{i,t}^T V_{i,zz(t-1)}^{-1} Z_{i,t})\ , \tag{16}$$

$$\lambda_{i,t} = \lambda_{i,t-1}\ \alpha_i^2\ \left(1 + (Y_t - \hat{P}_{i,t-1}^T Z_{i,t})^T \right.$$
$$\left.\lambda_{i,t-1}^{-1}\ (Y_t - \hat{P}_{i,t-1}^T Z_{i,t})(\alpha_i^2 + Z_{i,t}^T V_{i,zz(t-1)}^{-1} Z_{i,t})^{-1}\right)\ . \tag{17}$$

4 Border Detection

A pixel to be identified as a region border pixel is required to be confirmed in at least two different modelling directions in the same spectral band, i.e.

$$\omega_{m,n} = \begin{cases} \otimes & \text{if } \max_{t=(m,n,.)} \left\{ \sum_{i=\{\uparrow,\downarrow,\leftarrow,\rightarrow\}} \psi_{t^i} \right\} \geq 2 \\ \odot & \text{otherwise} \end{cases}$$

where $\otimes$ denotes the border class indicator and $\odot$ is the common indicator for texture classes present in the scene.

5 Numerical Realization

The predictors in (11) can be evaluated using updating of matrices $V_{i,t}$ (7) and their following inversion. Another possibility is the direct updating of $\hat{P}_{i,t}$ (9). To ensure the numerical stability of the solution, it is advantageous to calculate $\hat{P}_{i,t}$ (9) using a square-root filter, which guarantees the positivity of matrix (7). The filter updates directly the Cholesky factor of matrices $V_{i,t}^{-1}$.

Alternatively it is possible to use the UDU filter (a factorization into two triangular and one diagonal matrices) for this purpose. Initialisation of recursive (9) and (17) must keep the condition of positive definiteness of matrices $V_{i,0}$ (4). We implemented in our algorithm the uniform a priori start $V_{i,0} = I$. This solution not only conforms with the initial lack of information at the start of algorithm, but also simplifies the calculation of the integral (12). Another possibility could be for example a local condition start, which ensures a quicker adaptation .

6 Results

In this section we present segmentation results of the proposed method. Our goal was to properly detect single homogeneous textures present in the image. The performance criterion checks the texture area detected borders and compares them with the ground truth information. A version of our algorithm complemented with a region postprocessing step to remove undersegmentation introduced by undetected border pixels and noise border pixels will be reported elsewhere. The experimental texture mosaic 226×226 was created from five natural colour textures together with its gray scale alternative for the monospectral

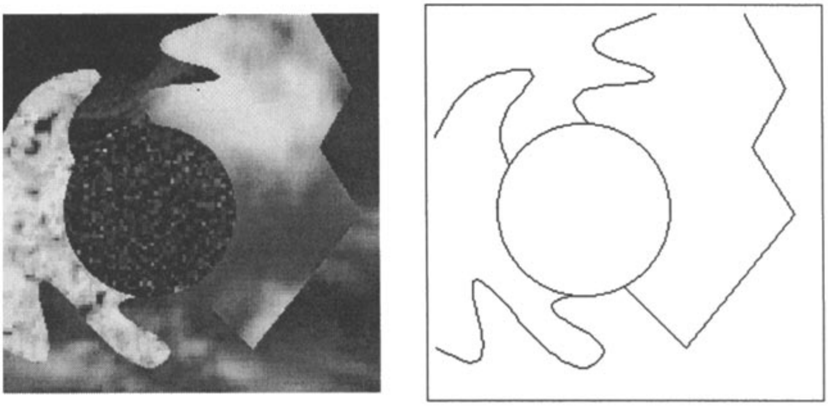

Fig. 1. Texture mosaic composed from five different textures and the optimal segmentation.

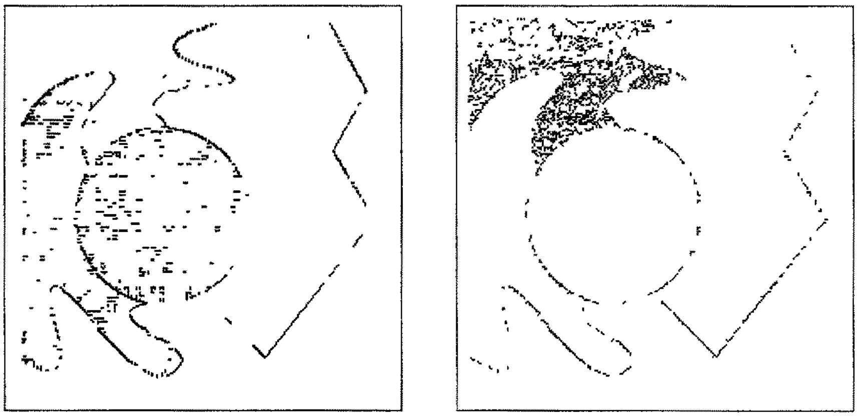

Fig. 2. Detected texture borders on the single-spectral and the colour image.

experiment. We have chosen natural textures rather than synthesized (for example using affine Markov random field models) because they are expected to be more difficult for the underlying segmentation model. The left image on Fig. 1 shows the gray scale version of the test image mosaic while the right image on Fig. 1 shows the ground truth texture regions borders. The results obtained on the gray scale image (Fig. 2 left) and on the colour mosaic (Fig. 2 right) are presented on the following images. The multispectral algorithm performs better on four from five mosaic components, however the left upper region has lot of false detections. These errors can be diminished with an adequate postprocessing.
Comparison of texture segmentation to the ground truth is done using the criteria of probability of finding a correct border pixel

$$P_c = \frac{card\{\mathcal{T} \cap \mathcal{G}\}}{card\{\mathcal{G}\}} \,, \qquad\qquad P_w = \frac{card\{\mathcal{T} \cap (\mathcal{I} - \mathcal{G})\}}{card\{\mathcal{I}\}}$$

and a probability of wrong border pixel detection where $\mathcal{G}, \mathcal{T}, \mathcal{I}$ are the ground truth set, the segmentation result set, and the set of all image pixels, respectively.

The optimal regression model M_i for the test image was found to be:

$$M_i \qquad \begin{array}{ccc} & * & \\ * & \circ & * \end{array}$$

where pixels corresponding to the contextual support set I_t are denoted $*$ and the predicted pixel $\circ$, respectively.

Table 1. Segmentation performance criteria

	gray scale test		colour test	
method	P_c	P_w	P_c	P_w
$\alpha = 0.99$	0.93	0.02	0.98	0.02
$\alpha = 0.91$	0.96	0.01	0.98	0.02
$\alpha = 0.80$	0.92	0.02	0.98	0.03

Results in Table 1. indicates the method dependence on the exponential forgetting factor α for the optimal model and the monospectral data while colour version seems to be more robust to the forgetting factor changes. These results shows good performance of our method even without additional region postprocessing step. The method is very fast, although we could not directly compare processing times of other unsupervised methods published we can estimate our processing time to be comparable to the quickest methods developed for texture segmentation. The presented method demonstrates low number of wrongly detected border pixels in most regions, but in the same time misses some correct border pixels. Detected borders are clean and accurately located.

7 Conclusion

We proposed the novel efficient and robust method based on a texture pixels prediction modelling. A texture is modelled using an adaptive causal regression model. The adaptive predictor uses spatial correlation from neighbouring data what results in improved robustness of the algorithm over rigid schemes, which are affected with outliers often present at the boundary of distinct textures. The proposed algorithm is recursive and therefore numerically effective. A parallel implementation of the algorithm is straightforward, every image row and column can be processed independently by its dedicated processor. The numerical stability is guaranteed using the Cholesky factorization of data gathering matrices. Usual handicap of segmentation methods is their lot of application dependent parameters to be experimentally estimated. Some methods need nearly a dozen adjustable parameters. Our method on the other hand requires only a contextual neighbourhood selection, which can be done using Bayesian statistics derived in the paper. The algorithm performance is demonstrated on the test natural texture mosaic, however more extensive testing is necessary. The preliminary test results of the algorithm are encouraging. The proposed method was always able to find all textures borders in our experiments with excellent border localization precision. However some postprocessing (gap filling, noise reduction) is still needed. The presented method segments textures independently in each spectral band and combines the majority voting results from all bands. An alternative can be a multidimensional regressive model with much larger set of parameters.

The proposed method is fully adaptive, numerically robust and still with moderate computation complexity so it can be used in an on-line virtual reality acquisition system, robot navigation system or some other image acquisition systems.

Acknowledgements

This research was supported by the INCO Copernicus project no. 960174 VIRTUOUS. The author thank his colleague Vojtěch Havlíček for preparing the experimental texture mosaic.

References

1. Broemeling,L.D.: Bayesian Analysis of Linear Models. New York, Dekker 1985
2. Haindl,M. " Texture synthesis ," *CWI Quarterly* 4 , 1991, no. 4, pp. 305-331.
3. Haindl,M., Šimberová, S.: A Multispectral Image Line Reconstruction method. In: Theory & Applications of Image Analysis. P. Johansen, S. Olsen Eds., World Scientific Publishing Co., Singapore, 1992
4. Haralick, R.M.: "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, 1979, pp. 786-804.
5. Ojala, T., Pietikainen, M.: "Unsupervised Texture Segmentation Using Feature Distributions," *Proc. of ICIAP'97*, Firenze, pp. 311-316, 1997.
6. Sinha, S.S., Jain, R.: Handbook of Pattern Recognition and Image Processing. Wiley, New York, 1994