A NEW METHOD FOR MULTI-RESOLUTION TEXTURE SEGMENTATION USING GAUSSIAN MARKOV RANDOM FIELDS

Roni Mittelman and Moshe Porat

Department of Electrical Engineering, Technion-Israel Institute of Technology
Haifa 32000, Israel
Phone: + (972) 48294684, Fax: + (972) 48295757, email: ronimi@tx.technion.ac.il, mp@ee.technion.ac.il
web: http://vision.technion.ac.il/mp

ABSTRACT

A new approach to multi-resolution modeling of images is introduced and applied to the task of semi-unsupervised texture segmentation using Gaussian Markov random fields (GMRFs). It is shown that traditional GMRF modeling of multi-resolution coefficients is incapable of accounting for the non-Gaussian statistics which often characterize the multi-resolution coefficients. On the other hand, the marginal distributions of the new approach can be closely modeled using a Gaussian distribution, and therefore lend itself efficiently to GMRF statistical modeling of images. Experimental results of texture segmentation using textures with non-Gaussian marginal distributions suggest that the new framework is superior to traditional GMRF modeling of the multi-resolution coefficients for segmentation of non Gaussian textures.

1. INTRODUCTION

There have been many applications presented in the literature in which Gaussian Markov random fields (GMRFs) were employed for the modeling of image statistics. GMRF modeling accounts for the correlation structure of an observation field while keeping the mathematical derivation tractable. A GMRF model assumes that an observation field is a realization of a multivariate Gaussian distribution with a certain parametric type of covariance matrix [1] which in turn (assuming a toroidal structure) necessitates the marginal distributions to be Gaussian. Therefore, GMRF modeling may not be efficiently utilized in cases where the marginal distribution of the observation field does not conform to a Gaussian distribution.

The GMRF model has been used to model various multi-resolution structured observation fields, where for the case of multi-resolution based texture segmentation, both the Gaussian pyramid and the wavelet transform have been employed [2,3]. When modeling the statistical structure of the wavelet coefficients of natural images, the GMRF model is often not applicable since the marginal distributions of the wavelet coefficients are characterized by a large peak at zero and heavier tails than a Gaussian of the same variance [4,5]. A similar argument applies to the Gaussian pyramid representation, which does not conform to a Gaussian marginal distribution. Early computer vision algorithms ignored the non-Gaussian statistics and used features of the first two moments exclusively [6]. To account for these marginal statistics, the generalized Gaussian parametric distribution has often been used, however it cannot account for the statistical dependencies between adjacent coefficients. Another approach which accounts for the marginal statistics without the assumption of statistical independence of the wavelet coefficients is the so called hidden Markov trees [7], where each coefficient was assumed a realization of a Gaussian mixture distribution, where the statistical dependence between the coefficients is modeled using a tree structured hidden Markov model. In a previous work [8], we have introduced new feature statistics, which we applied to wavelet-based texture classification and showed classification results superior to [9]. The feature is based on the local second moment estimates of the wavelet coefficients which distribution was shown to be logarithmically scaled. This work complements [8], by applying the feature to GMRF based semi-unsupervised texture segmentation.

2. LOCAL SECOND MOMENT ESTIMATES

2.1 The Gaussian scale mixtures (GSM) distribution

The marginal distributions of the wavelet coefficients of natural images are characterized by a large pick at zero and heavier tails than a Gaussian of the same variance. There have been a number of parametric distribution models suggested that can account for these non-Gaussian statistics, such as the generalized Gaussian, or Gaussian mixture distributions. In [4,5] GSM were proposed as a distribution model that can account for both the marginal and joint distribution properties of local neighborhoods of the coefficients.

The probability density of a GSM variable is of the form:

$$p(X) = \int p(X \mid z)p(z)dz \quad , \quad p(X \mid z) \sim N(0, z \cdot C_u)$$

C_u is a covariance matrix, and z is a positive scalar random variable referred to as the hidden multiplier, and X is a random vector composed of the neighborhood's coefficients.

2.2 Maximum likelihood estimator of the hidden multipliers

Assuming that the covariance matrix is of the form $C_u = \sigma_u^2 I$ (this is valid for the wavelet transform since the wavelet coefficients are roughly de-correlated) and setting $\sigma_u^2 = 1$ (ignoring the proportionality constant σ_u^2), the maximum likelihood estimator is given by:

$$(1) \quad \hat{z}(X) = \frac{1}{N} X^T X \,,$$

where N is the number of coefficients within the neighborhood. Therefore the hidden multiplier's estimate corresponds to the local second moment estimate.

2.3 A parametric distribution for the logarithm of the local second moment estimates

The distribution of the hidden multipliers in natural images, was shown empirically in [5] to be closely modeled by a Gaussian density on a logarithmic scale. Therefore we expect the logarithm of the local second moment estimates to follow the same density model. Since the hidden multipliers' distribution in [5] was derived directly from the dataset without the intermediate stage of computing the local second moment estimates, we examined the empirical distribution of the logarithm of the local second moment estimates. Fig. 1 shows the average over the histograms of 180 sub-bands of 20 texture images where each histogram was normalized for mean and variance. It can be verified that the density of the logarithm of the local second moment estimates of the wavelet coefficients, closely follows a Gaussian distribution.

2.4 A new feature space

We compute the logarithm of the local second moment estimate for each overlapping 3x3 neighborhood in each subband of the wavelet transform using (1), where we assumed a toroidal structure for the boundaries. Thus a new feature space is obtained, where the marginal distribution of the features in each sub-band is closely Gaussian. Since the logarithm of the local second moment estimates of the wavelet coefficients is computed for overlapping neighborhoods, there is a large spatial correlation among neighboring observations in the new features. Since GMRF modeling is equivalent to assuming a multivariate Gaussian distribution with a certain parametric type of covariance matrix [1] (assuming a toroidal structure) it lends itself efficiently to statistical modeling of the new feature space. Although the cooccurrence histograms of the local second moment estimates which we have examined did not conform to a multivariate Gaussian distribution, we believe that since the marginal distribution is Gaussian, modeling the new feature space using a GMRF is more accurate than modelling other feature spaces such as the Gaussian pyramid or the wavelet transform using a GMRF.

3. SEMI-UNSUPERVISED TEXTURE SEGMENTATION USING GMRF

Our texture segmentation algorithm follows [2] with a few major differences described in section 3.4. We consider semi-unsupervised segmentation where the number of texture classes is known in advance, and the parameters of the observation field are estimated in an unsupervised manner. The parameters for the label field were selected experimentally.

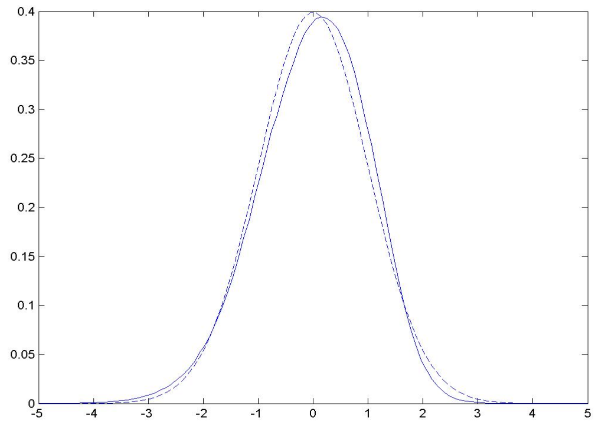

Fig. 1. Empirically measured histograms for the logarithm of the local second moment estimates compared to a Gaussian density (dashed line).

3.1 Statistical models

In order to perform texture segmentation, a doubly stochastic model is used, where Y is the observed field and X is the class label field, which is defined on a multi-resolution lattice. We denote $X^{(n)}$ to be a raster scanned vector of the class labels assigned to level n of the lattice, which can take values from the class label set $\{1,\ldots,M\}$. Its distribution follows a multi-scale Markov random field, where both intra-scale and inter-scale pair-wise clique types, denoted C_1 and C_2 respectively, are used. C_1 is the set of all first order neighborhoods in the same level [1], and C_2 is the set of all the neighborhoods which include a node at level n, and its parent at level $n+1$, and its four sons at level $n-1$ (level L is the coarsest level and 1 is the finest level).
Accordingly, the probability mass function for X is given by:

$$p_X(x) = \frac{1}{z}\exp\{-\sum_{r,s\in C_1} \beta_1^{(n(s))}(1-\delta(x_r - x_s)) - \sum_{r,s\in C_2} \beta_2^{(n(s))}(1-\delta(x_r - x_s))\},$$

where $n(s)$ is the level of node s, and z here is a normalizing constant.
We denote $Y^{(n,k)}$ to be a raster scanned vector of the observed data field at level n which is assumed to be a causal GMRF model conditioned on the class label field, where the number of prediction coefficients is R. We assume that each of the multi-resolution's scales and orientations are statistically independent. Therefore the conditional probability density function of Y given X is

$$f_{Y|X}(y \mid x) = \prod_{n=1}^{L}\prod_{k=1}^{O(n)}\prod_{s\in S(n,k)} \frac{1}{\sqrt{2\pi\sigma_s^2}} \exp\{-\frac{(\tilde{y}_s^{(n,k)})^2}{2\sigma_s^2}\}$$

$$\tilde{y}_s^{(n,k)} = y_s^{(n,k)} - \mu^{(n,k)} + \sum_{r>0} a_r^{(n,k)}(y_{s-r}^{(n,k)} - \mu^{(n,k)}),$$

where $S(n,k)$ is the set of all observation nodes in level n and orientation k, and $O(n)$ is the number of orientations in level n, where we define the coarsest scaling sub-band to be the fourth orientation in level L.

3.2 Texture segmentation

The optimization criteria used to perform the segmentation is the minimization of the expected value of the misclassified nodes in the multi-resolution lattice. The decision rule that minimizes this criterion can be shown to be given by the maximum of the posterior marginals (MPM):

$$X_s = \underset{c}{\operatorname{argmax}}\; p_{X_s|Y}(c \mid y).$$

The posterior marginals can be approximated as follows. First, a Markov chain $X(t)$ is generated using a Gibbs sampler with constant temperature, which converges to a random field with probability mass function $p_{X|Y}(x \mid y,\theta)$ where θ is the set of the GMRF parameters, and where:

$$p_{X|Y}(x \mid y,\theta) = \frac{f_{Y|X}(y \mid x,\theta)\, p_X(x)}{f_Y(y \mid \theta)}.$$

Subsequently, the posterior marginals are computed as the fraction of time the Markov chain spends in each label at each node of the lattice:

$$p_{X_s|Y}(x_s = c \mid y) = \frac{1}{T}\sum_{t=1}^{T} \delta(x_s(t) - c),$$

where T is the number of iterations used in the Gibbs sampler.

3.3 Parameter estimation

The parameter estimation for the GMRF model of each texture class is performed using an EM like algorithm where the segmentation and parameter estimation stages are repeated alternately. In each iteration, first the posterior marginals are computed using the current parameter estimates and the Gibbs sampler, as has been described previously. Subsequently the parameters are updated using the following equations (These equations are obtained by using equation (2) for the class means, and maximizing the EM Q-function for $\hat{a}_{c,q}^{(n,k)}$, $\sigma_c^{(n,k)}$ [2]):

$$\forall n = 1 \ldots L,\ \forall k = 1 \ldots O(n),\ \forall c = 1 \ldots M:$$

$$(2)\quad \hat{\mu}_c^{(n,k)} = \frac{1}{N_c^{(n,k)}} \sum_{s\in S(n,k)} y_s\, p(x_s = c \mid Y = y, \theta_{old}),$$

$$\text{where } N_c^{(n,k)} = \sum_{s\in S(n,k)} p(x_s = c \mid Y = y, \theta_{old}).$$

$$\forall n = 1 \ldots L,\ \forall k = 1 \ldots O(n),\ \forall c = 1 \ldots M,\ \forall r = 1 \ldots R:$$

$$(3)\quad \sum_{s\in S(n,k)} [y_s - \hat{\mu}_c^{(n,k)} + \sum_{q>0} \hat{a}_{c,q}^{(n,k)}(y_{s-q} - \hat{\mu}_c^{(n,k)})] \cdot (y_{s-r} - \hat{\mu}_c^{(n,k)})\, p(x_s = c \mid Y = y, \theta_{old}) = 0\ .$$

$$(4)\quad \sum_{s\in S(n,k)} [y_s - \hat{\mu}_c^{(n,k)} + \sum_{q>0} \hat{a}_{c,q}^{(n,k)}(y_{s-q} - \hat{\mu}_c^{(n,k)})]^2 \cdot p(x_s = c \mid Y = y, \theta_{old}) - (\sigma_c^{(n,k)})^2 N_c^{(n,k)} = 0,$$

where θ_{old} is the set of old parameters estimates.

The class means are first computed using (2), subsequently the class prediction coefficients are computed using (3), and finally the class variances are computed using (4).

3.4 Comparison to Comer and Delp [2]

Our texture segmentation algorithm differs from [2] in three major details:

1) The observation field in [2] was the Gaussian pyramid representation, therefore it employed only one orientation as opposed to three in our method.
2) Since the filter coefficients used to obtain the Gaussian pyramid representation in [2] sum to one, the class means did not depend on the resolution level, whereas we use different class means for each resolution level.
3) The different levels in the Gaussian pyramid in [2] were not assumed to be statistically independent. The pyramid representation was divided into quad-trees, where each quad-tree was scanned from the coarsest resolution to the finest by a predefined raster scan order. Our approach on the other hand assumes each of the sub-bands to be statistically independent, therefore the observations of each sub-band are raster scanned, ignoring all other levels and orientations.

4. EXPERIMENTAL RESULTS

In order to test the new feature space for multi-resolution based segmentation of textured images, we implemented the texture segmentation algorithm described in Section 3, where the new feature was used. We compared the results obtained using the new feature both to those obtained using the segmentation algorithm described in [2], where the Gaussian pyramid was used as the observation data, and to those obtained when using the wavelet transform with the texture segmentation algorithm described here.

4.1 Implementation details

We used three levels both for the wavelet transform and four levels for the Gaussian pyramid. The values of β_1 and β_2 were identical for each of the four levels where we used $\beta_1 = 1.6, \beta_2 = 0.8$, and three prediction coefficients were used for the causal GMRF model. In each iteration of the parameter estimation stage we randomly initialized the label field (but for the last iteration). The segmentation result is that of the finest resolution level.

4.2 Segmentation results

The experiments in Fig. 2 show that all the algorithms successfully separated the two textures since the two textures are very different. The results in Fig. 3 show a significant improvement when using our texture segmentation method, both when using the wavelet transform and when using the new feature, compared to the method described in [2]. Fig. 4 shows a significant improvement when using the new feature compared to using the wavelet transform, which shows that the Gaussian statistics of the marginal distributions of the new feature are of major importance for GMRF statistical modeling.

5. CONCLUSIONS

We have presented a new feature space that can be used for multi-resolution based statistical modeling of images. Since the marginal distributions of the sub-bands of the new feature are closely Gaussian, it facilitates the statistical description of the joint distribution, compared to other pyramid representations such as the Gaussian pyramid, or the wavelet transform, which show non-Gaussian statistics. Our main conclusion is that the new feature is more appropriate to texture segmentation using GMRF compared to other representations such as the Gaussian pyramid or the wavelet transform.

REFERENCES

[1] S. Krishnamachari and R. Chellapa, "Multiresolution Gauss-Markov random field models for texture segmentation", *IEEE Trans. On Image Proc.*, vol. 6 no. 2, pp. 251-267, 1997.

[2] M. L. Comer and E. J. Delp, "Segmentation of textured images using a multiresolution Gaussian autoregressive model", *IEEE Trans. On Image Proc.*, vol. 8, no. 3, 408-420, 1999.

[3] H. Noda, M. N. Shirazi and E. Kawaguchi, "Textured image segmentation using MRF in wavelet domain", *ICIP* 2000, 572-575.

[4] V. Strela, J. Portilla, and E. Simoncelli, "Image denoising using a local Gaussian scale mixture of Gaussians in the wavelet domain", *SPIE* 2000, pp. 363-371.

[5] J. Portilla, V. Strela, M. J. Wainwright and E. P. Simoncelli, "Adaptive Wiener denoising using a Gaussian scale mixture model in the wavelet domain", *ICIP* 2001, pp. 37-40.

[6] M. Porat, and Y. Zeevi, "Localized Texture Processing in Vision: Analysis and synthesis in the Gaborian space," *IEEE Trans. On Biomedical Engineering*, Vol. 36, No 1, pp. 115-129, 1989.

[7] M. S. Crouse, R. D. Nowak, and R. G. Barniuk, "Wavelet-based statistical signal processing using hidden Markov models", *IEEE Trans. On Signal Proc.*, Vol. 46, No. 4, pp. 886-902, 1998.

[8] R. Mittelman and M. Porat, "A new approach to feature extraction for wavelet-based texture classification", to be presented in *ICIP* 2005.

[9] X. Liu, and D. Wang, "Texture classification using spectral histograms", *IEEE Trans. On Image Proc.*, Vol. 12, No. 6, pp. 661-670, 2003.

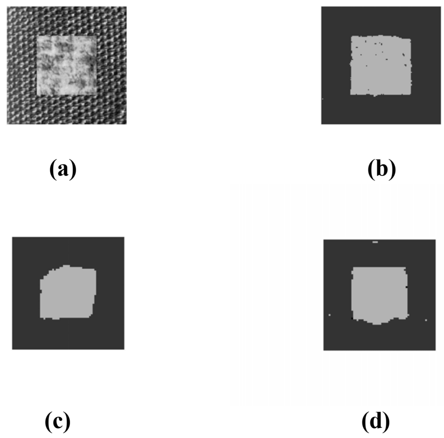

Fig. 2. **Segmentation results: (a) Original image (size 128x128), (b) Segmentation results using the Gaussian pyramid as described in [2], (c) Segmentation results using the wavelet transform, (d) Segmentation results using new feature space.**

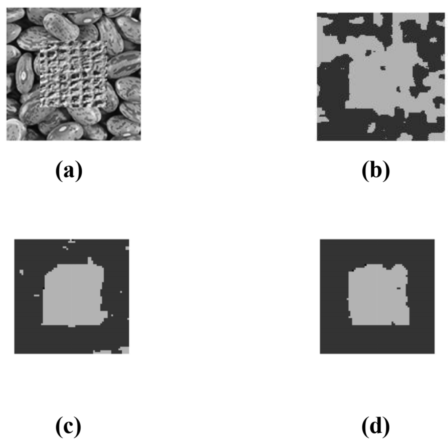

Fig. 3. **Segmentation results: (a) Original image (size 128x128), (b) Segmentation results using the Gaussian pyramid as described in [2], (c) Segmentation results using the wavelet transform, (d) Segmentation results using new feature space.**

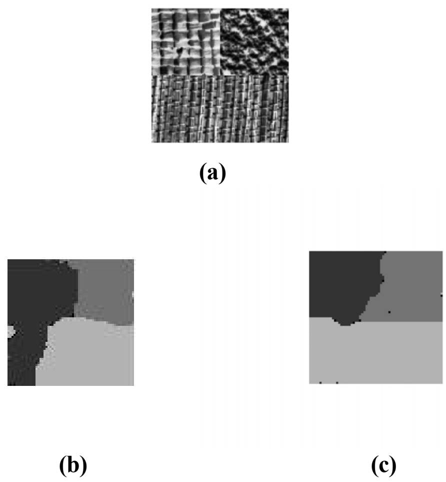

Fig. 4. **Segmentation results: (a) Original image (size 128x128), (b) Segmentation results using the wavelet pyramid, (c) segmentation results using the new feature space.**