Color Image Classification and Parameter Estimation in a Markovian Framework

Zoltan Kato, Ting-Chuen Pong, John Chung-Mong Lee
Hong Kong University of Science and Technology
Computer Science Department, Clear Water Bay, Kowloon, Hong Kong
Tel: +852 2358 7000 — Fax:+852 2358 1477
email: `kato@cs.ust.hk`, `tcpong@cs.ust.hk`, `cmlee@cs.ust.hk`

Abstract

In this paper, we propose an unsupervised color image classification algorithm based on a Markov random field (MRF) model. In the MRF model, we use the CIE-**luv** color metric because it is close to human perception when computing color differences. On the other hand, intensity and chroma information is separated in this space. Without parameter estimation, our model would not be useful in real-life applications. We propose herein a new method to estimate mean vectors effectively even if the observed image is very noisy and the histogram does not have clearly distinguishable peaks. These values are then used in a more complex, iterative estimation process as initial values. The only parameter supplied by the user is the number of classes. All other parameters are estimated from the observed image. The algorithm has been tested on a variety of real images (indoor, outdoor), noisy video sequences and noisy synthetic images.

1 Introduction

Image classification is an important early vision task where pixels with similar features are grouped into homogeneous regions. Many high level processing tasks (surface description, object recognition, for example) are based on such a preprocessed image. Using color information can considerably improve capabilities of image classification algorithms compared to purely intensity-based approaches.

Some examples of MRF color image models can be found in [2, 9, 11]. In [2], a MRF segmentation model is proposed and the use of three different lattice schemes (squares, hexagons and triangles) are discussed. A multi-resolution approach is presented in [9]. A color texture MRF model is proposed in [11].

Usually, MRF-based classification methods suffer from a lack of parameter estimation. The majority of the proposed methods are supervised, which limits their practical use because human intervention is needed to compute the model parameters. Herein, we are interested in *data driven* algorithms since in real life applications model parameters are usually unknown, one has to estimate them without human intervention. Estimation algorithms are usually iterative [10, 7], subsequently generating a labeling, estimating parameters from it, then generating a new labeling using these parameters, etc ... For such a method, we need a reasonably good initial value for each parameter. We propose here a new method to find the components of a color histogram taking into account spatial information. The basic idea is to requantize the observed image via a pre-segmentation. Using this algorithm, we develop an unsupervised color image classification algorithm. The only parameter supplied by the user is the number of classes.

2 A Color Image Classification Model

We use the CIE-**luv** [5] color space herein because it separates luminance and chroma information and it is close to human perception when computing color differences. The model proposed hereafter is based on our earlier work on intensity based classification [6]. Let us suppose that the observed image consists of three spectral component values (**luv**) at each pixel denoted by the vector $\vec{f_s}$. We are looking for the labeling $\hat{\omega}$ which maximizes the a posteriori probability $P(\omega \mid \mathcal{F})$, that is the *maximum a posteriori* (MAP) estimate:

$$\hat{\omega} = \arg\max_{\omega \in \Omega} \prod_{s \in \mathcal{S}} P(\vec{f_s} \mid \omega_s) \prod_{C \in \mathcal{C}} \exp(-V_C(\omega_C)) \ . \quad (1)$$

A natural assumption is that $P(\vec{f_s} \mid \omega_s)$ is Gaussian, the classes $\lambda \in \Lambda = \{1, 2, \ldots, L\}$ are represented by the mean vectors $\vec{\mu}_\lambda$ and the covariance matrices Σ_λ.

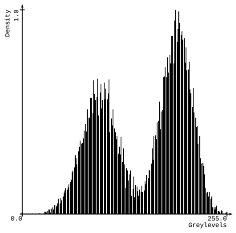

Figure 1: Histogram of a noisy image's **v** component.

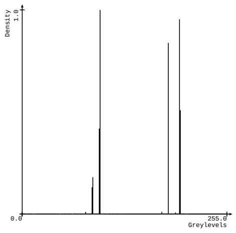

Figure 2: Histogram of the noisy image's **v** component after pre-segmentation.

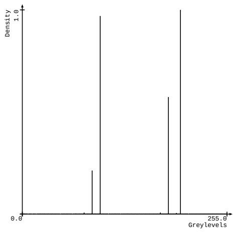

Figure 3: Histogram of the noisy image's **v** component after pre-segmentation and quantization.

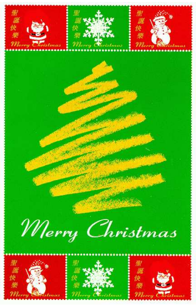

Figure 4: Original postcard image

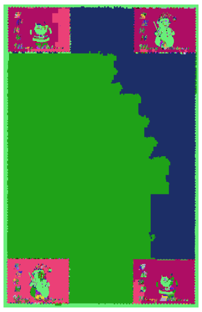

Figure 5: Result after pre-segmentation and quantization colored by random colors

Figure 6: Final unsupervised classification colored by random colors

class	$\mu(\mathbf{l})$	$\mu(\mathbf{u})$	$\mu(\mathbf{v})$	$\Sigma(\mathbf{l},\mathbf{l})$	$\Sigma(\mathbf{u},\mathbf{u})$	$\Sigma(\mathbf{v},\mathbf{v})$
Initial values						
1	—	97	87	1	1	1
2	—	82	97	1	1	1
3	—	42	197	1	1	1
4	—	47	182	1	1	1
Final estimates						
1	100	97	86	513	203	456
2	97	80	98	464	202	486
3	189	41	197	280	137	312
4	132	49	181	365	177	410
Supervised values						
1	105	98	87	203	89	210
2	97	81	99	465	208	522
3	189	41	197	82	41	98
4	132	48	182	375	188	429

Table 1: Parameter values of the noisy synthetic image

It is then clear that the energy function of the so defined MRF image model has the following form:

$$
\begin{aligned}
U(\omega, \mathcal{F}) &= \sum_{s \in \mathcal{S}} \Big(\ln(\sqrt{(2\pi)^3 \mid \Sigma_{\omega_s} \mid}) \\
&+ \frac{1}{2}(\vec{f}_s - \vec{\mu}_{\omega_s})\Sigma_{\omega_s}^{-1}(\vec{f}_s - \vec{\mu}_{\omega_s})^T \Big) \\
&+ \beta \sum_{\{s,r\} \in \mathcal{C}} \delta(\omega_s, \omega_r),
\end{aligned} \tag{2}
$$

$$
\delta(\omega_s, \omega_r) = \begin{cases} 0 & \text{if } \omega_s = \omega_r \\ 1 & \text{if } \omega_s \neq \omega_r \end{cases} \tag{3}
$$

where $\beta > 0$ is a hyper-parameter controlling the homogeneity of the regions. As β increases, the resulting regions become more homogeneous.

3 Parameter Estimation

Our goal is to propose an unsupervised classification algorithm. Therefore, we need a method to estimate the mean vector $\vec{\mu}_\lambda$ and the covariance matrix Σ_λ for each class, and the hyper-parameter β, which controls the homogeneity of the regions. Since we do not have a labeled data set, we cannot use classical estimation methods such as Maximum Likelihood (ML). We have to generate labeled samples from the observed

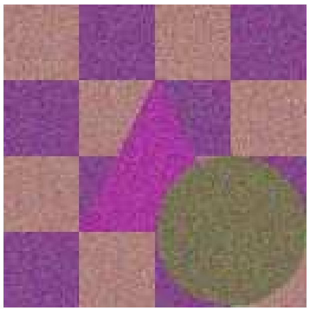

Figure 7: Noisy synthetic image (SNR=5dB)

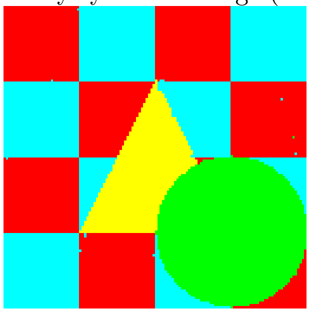

Figure 8: Supervised classification result using color information.

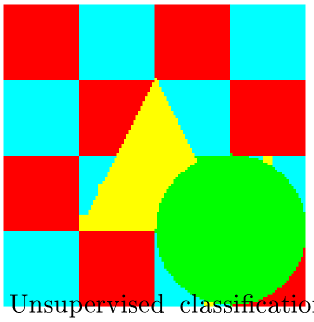

Figure 9: Unsupervised classification result using color information.

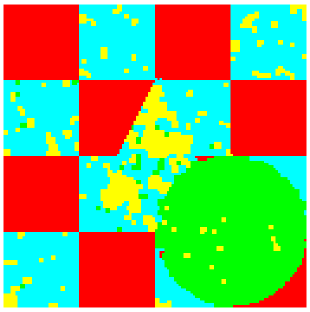

Figure 10: Supervised classification result using only intensity information.

Figure 11: Original outdoor image (photo by Eva Kisgyorgy).

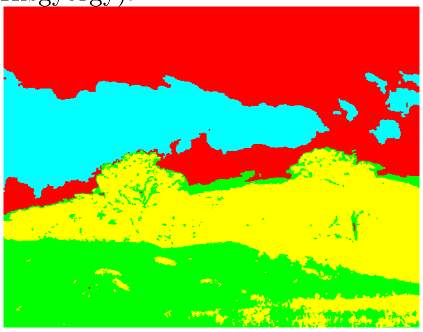

Figure 12: Unsupervised classification result.

Figure 13: Original image extracted from a noisy video sequence.

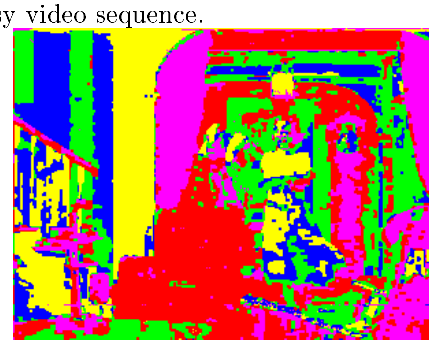

Figure 14: Unsupervised classification result.

image in order to be able to use the ML estimator. In statistics, the problem is known as the *incomplete data* problem. A broadly applicable algorithm has been proposed by Dempster *et al.* [3], called *Expectation – Maximization* (EM). The algorithm aims at determining the ML estimate of the parameters Θ by making use of the estimation of the missing data (i.e. the label field).

Herein, we use an EM-like adaptive classification algorithm [4, 8] doing classification and parameter estimation simultaneously (for more details, see [7]:

Algorithm 1 (Adaptive Classification)

① *Set $k = 0$ and initialize $\widehat{\Theta}^0$.*

② *Maximize $P(\omega \mid \mathcal{F}, \widehat{\Theta}^k)$ using an optimization algorithm (ICM [1], for instance). The resulting labeling is denoted by $\widehat{\omega}^{k+1}$.*

③ *Update the current estimate of the parameters, $\widehat{\Theta}^{k+1}$ to the ML estimate based on the current labeling $\widehat{\omega}^{k+1}$.*

④ *Goto Step ② with $k = k + 1$ until $\widehat{\Theta}^k$ stabilizes.*

4 Obtaining Initial Parameters

The estimation procedure described previously supposes that we have an initial guess about the parameters. The most crucial is the mean value, all the other parameters are far less sensitive to initialization. Estimating the mean values is a classical problem, namely the determination of the components of a Gaussian mixture without any *a priori* information. Unfortunately, classical methods [12] will fail if the histogram does not have clearly distinguishable peaks, which is often the case in dealing with noisy images. For example, in Figure 1, we show the histogram of a noisy synthetic image (SNR = 5dB). Any histogram-based method will fail to find a reasonably good mean value for the 4 classes we have on this image. However, using our segmentation-based initialization method, we are able to obtain the histogram shown in Figure 3 which has four clearly distinguishable peaks.

Another problem, specific to color images is a sparse histogram. Consider a color image with 24 bit color codes. Clearly, we have $256^3 = 16777216$ possible colors, which is usually much greater than the number of pixels in an image (even a 1024×1024 image has only 1048576 pixels). As a result, the histogram of such an image will be too sparse for statistical analysis. Typically, less than 10 pixels belong to the same color value yielding a completely flat histogram.

To solve this problem, we have to re-quantize our image in order to reduce the number of possible colors. However, using a classical quantization algorithm would transform the original image into a coarse one losing important spatial information necessary for the classification. Thus, we have to re-quantize the image taking into account spatial information.

Now, let us see our approach. It is based on a pre-segmentation instead of analyzing the histogram of the observed image. The initial segmentation is obtained via a split and merge algorithm using color difference as a homogeneity measure. Thanks to the **luv** color space, color difference is easily obtained as the second norm of two color vectors. Regions are represented by the mean vector of the original pixels. Two neighboring regions are merged if their color difference is less than a certain threshold τ. A smaller τ results in smaller but more regions, a larger τ gives larger but fewer regions. Obviously, to keep the method unsupervised, we have to determine τ from the observed image. This is easily achieved because we only need a *reasonably* good segmentation, thus the number or the size of regions is not crucial. In practice, we have found that τ may be obtained as a certain percent of the maximal color difference in the observed image (in our tests, we have used 20%).

The next step is to quantize the obtained image. This is needed because we often obtain regions which are not neighbors but their colors are very close (see Figure 2). This might cause detection of false peaks and lead to wrong initial estimations. Experiments show that a 20% reduction of the number of gray-levels is sufficient (see Figure 3).

From the histogram obtained by the pre-segmentation and quantization process, we can easily extract the peaks according to the number of classes (the only parameter known *a priori*). At this stage, we only use the **u** and **v** components because they carry chromatic the information and the histogram of the **uv** space is denser and thus peaks are easier to detect.

It may seem strange to use a segmentation algorithm to get initial parameters for a pixel classification algorithm, which is very similar to a segmentation. Why not use the result of the pre-segmentation? To clarify this issue, consider the image in Figure 4. It contains large homogeneous regions but also small, fine details. Since the pre-segmentation algorithm uses only the color difference (because we do not have other information at this stage), it is fast but unable to keep all the fine details (Figure 5). But this result is perfectly usable to re-quantize the image so that important spatial information is kept. The final result is then obtained by the more elaborate MRF model (Figure 6).

5 Experimental Results

The proposed algorithm has been tested on a variety of images including synthetic noisy images, outdoor and indoor scenes, and video sequences. Herein, we present a few of these results and also compare color- and intensity-based classification, supervised and unsupervised results.

In Figure 8, we show the results obtained from a 128×128 noisy color synthetic image (see Figure 7) by *supervised* classification. The signal to noise ratio (SNR) was 5dB in this case. In Figure 10, we show the corresponding *supervised* classification using only intensity information. It can be seen that using color information can improve significantly the final result. The findings were the same for real images.

Let us see now the comparison between supervised and unsupervised classification. Figure 9 shows the result obtained by unsupervised classification of the noisy image with SNR=5dB and Table 1 shows the parameters. Of course, the supervised method performs better, especially when SNR increases. However, up to a certain SNR, the unsupervised results are very close to the supervised ones. For this particular image, this limit seems to be around SNR=5dB. The big advantage of the unsupervised method is that it *does not* requires human intervention and is fully data-driven.

Finally, a real scene segmentation is presented in Figure 12 (596×458). Figure 14 (304×228) shows the result obtained for an image extracted from a noisy video sequence. The computer time on a Sparc Station 10 was around 100 minutes for the largest image including the pre-segmentation and quantization needed for parameter initialization.

6 Conclusion

In this paper, we have proposed an unsupervised color image classification algorithm. The classification model is defined in a Markovian framework and uses a first order potential derived from a three-variate Gaussian distribution in order to tie the final classification to the observed image. To estimate the model-parameters, we have to use an iterative algorithm, which subsequently generates a labeling and then recomputes the parameter values. This process requires a good initial value for the parameters, especially for the mean values. Due to the large number of possible colors, the histogram cannot be analyzed directly, we have to re-quantize it without losing any important spatial information. To solve this problem, we have proposed a new method. It uses a pre-segmentation step based on color differences in order to reduce the number of colors. The method has been tested on a variety of real and synthetic images and the results are very close to supervised ones if SNR is kept reasonably high.

References

[1] J. Besag. On the statistical analysis of dirty pictures. *Jl. Roy. Statis. Soc. B.*, 1986.

[2] M. J. Daily. Color Image Segmentation Using Markov Random Fields. In *Proc. DARPA Image Understanding*, 1989.

[3] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *J. Roy. Statist. Soc., ser. B*, vol. 39(1):1–38, 1977.

[4] D. Geman. Bayesian Image Analysis by Adaptive Annealing. In *Proc. IGARSS'85*, pages 269–277, Amherst, USA, Oct. 1985.

[5] A. K. Jain. *Fundamentals of Digital Image Processing.* Prentice Hall, 1989.

[6] Z. Kato, M. Berthod, and J. Zerubia. A Hierarchical Markov Random Field Model and Multi-Temperature Annealing for Parallel Image Classification. *IEEE-CVGIP Graphical Models and Image Processing*, 58(1):18–37, Jan. 1996.

[7] Z. Kato, M. Berthod, J. Zerubia, and W. Pieczynski. Unsupervised Adaptive Image Segmentation. In *ICASSP'95*, Detroit, USA, May 1995.

[8] S. Lakshmanan and H. Derin. Simultaneous Parameter Estimation and Segmentation of Gibbs Random Fields Using Simulated Annealing. *IEEE–PAMI*, 11(8):799–813, Aug. 1989.

[9] J. Liu and Y. H. Yang. Multiresolution Color Image Segmentation. *IEEE-PAMI*, 16(7):689–700, July 1994.

[10] P. Masson and W. Pieczynski. SEM Algorithm and Unsupervised Statistical Segmentation of Satellite Images. *IEEE Geoscience and Remote Sensing*, 31(3):618–633, May 1993.

[11] D. K. Panjwani and G. Healey. Markov Random Field Models for Unsupervised Segmentation of Textured Color Images. *IEEE - PAMI*, 17(10):939–954, Oct. 1995.

[12] D. M. Titterington, A. F. M. Smith, and U. E. Makov. *Statistical Analysis of Finite Mixture Distributions.* John Wiley & Sons, 1985.