Texture Segmentation Comparison Using Grey Level Co-occurrence Probabilities and Markov Random Fields

David A. Clausi and Bing Yue
Department of Systems Design Engineering
University of Waterloo
Waterloo, Ontario, Canada N2L 3G1
dclausi@engmail.uwaterloo.ca, Bing.Yue@CCRS.NRCan.gc.ca

Abstract

The discrimination ability of texture features derived from Gaussian Markov random fields (*GMRFs*) and grey level co-occurrence probabilities (*GLCPs*) are compared and contrasted. More specifically, the role of window size in feature consistency and separability as well as the role of multiple textures within a window are investigated. GLCPs are demonstrated to have improved discrimination ability relative to MRFs with decreasing window size, an important concept when performing image segmentation. On the other hand, *GLCPs* are more sensitive to texture boundary confusion than *GMRFs*.

1 Introduction

Texture, a representation of the spatial relationship of grey levels in an image, is an important characteristic for computer image interpretation. Many texture feature methods exist [7], however, limited research has been conducted to compare different methods. Comparison texture papers often only consider the supervised classification problem, without considering full image segmentation [3]. In the case of unsupervised segmentation, windows contain mixed classes with unknown parameters, making the feature extraction and class assignment decisions far more challenging.

This paper compares the unsupervised segmentation capabilities of two popular texture methods: *GLCPs* (grey level co-occurrence probabilities) and *GMRFs* (Gaussian Markov random fields). Most notably, the paper emphasizes the role of window size selection when using *GLCP* and *GMRF* texture features for unsupervised segmentation. Test data includes *MRF* generated, Brodatz, and synthetic aperture radar (SAR) sea ice imagery.

2 Texture Feature Methods

GLCPs represent the conditional joint probabilities of all pairwise combinations of grey levels (i,j) in the fixed-size spatial window given interpixel distance (δ) and orientation (θ) [6]. Here, θ = 0, 45, 90, 135 degrees and $\delta = 1$ are used. To generate texture features, statistics (dissimilarity, entropy, and correlation [4]) are applied to the probabilities. The grey level quantization level is fixed at 16. For simplicity, k-means [5] is used to perform *GLCP* segmentation.

MRFs are recognized for being effective for texture analysis [2]. The basic premise is that neighborhood pixels are expected to have similar characteristics. Under the assumption of a Gaussian MRF (or GMRF), the following model is produced [2]:

$$x_s = \sum_{r \in \ N_s} \theta_r (x_{s+r} + x_{s-r}) + e_s \tag{1}$$

where x_s is a real number representing the center pixel of the neighborhood, x_{s+r} and x_{s-r} are a pair of pixels centered around x_s, e_s is a zero mean Gaussian noise, and θ_r represents the MRF model parameters. The summation is over some neighborhood N_s, as defined by the model order. The MRF parameters are determined using least squares. The iterated conditional mode (ICM) method is used for *GMRF* image segmentation [1].

Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04)
1051-4651/04 $ 20.00 IEEE

3 Methods and Testing

Large windows produce better estimates of texture, however, they can also lead to the undesirable situation of containing multiple texture classes. Small windows are less likely to contain multiple classes, however, the limited coverage can produce misleading features.

Research Question One *How does window size influence the estimated individual GLCP texture features and GMRF model parameters?*

Theory: For both *GLCP* and *GMRF*, different n generate different feature estimates. The effect of n on the stability of *GLCP* texture features and *GMRF* model parameters must be assessed.

Method: For a given texture, the relative change of a feature's standard deviation as a function of n is calculated. One MRF synthetic texture (1024 $\times$ 1024), one Brodatz texture (1024 $\times$ 1024) (pigskin), and one RADARSAT SAR texture (768 $\times$ 768) are used. From each texture image, 60 window samples are randomly selected for each of $n = 8, 16, 32, 64, 96$. The standard deviation σ per feature per window size is determined for each set of 60 samples. To measure the relative change, σ is normalized by σ for $n = 96$.

Results: Table 1 summarizes the average increase of σ across n. First, with decreasing n, estimated *GLCP* features and the *GMRF* model parameters increase exponentially. Second, with decreasing n, the σ of each *GMRF* model parameter increases faster than the *GLCP* features. The *GMRF* method requires a relatively larger n than the *GLCP* to obtain the same degree of stability in the feature estimates.

Research Question Two *How does n influence the cluster separability of the estimated features?*

Theory: This research question compares *GLCP* texture features versus *GMRF* model parameters for feature space separability. If the feature space separability is larger, it is assumed that those features are more appropriate for classification.

Method: The Fisher criterion [5] (J) is used as a non-parametric measure of the cluster separability. Further insight can be obtained by calculating the upper bound of classification error between feature cluster pairs using the Bhattacharyya error bound [5] (BEB). Three texture pairs are used for testing. These include an MRF generated synthetic image (with two textures) (1024 $\times$ 1024), a Brodatz image (1024 $\times$ 1024) containing wood grain and raffia and a SAR sea ice image (768 $\times$ 768) containing first year and multiyear ice. Sixty sample windows with sizes 8, 16, 32 and 64 are randomly selected from each texture in each image.

Table 1: Ratio of the standard deviations for each window size $(64, 32, 16, 8)$ with respect to window size 96 for both *GLCP* and *GMRF* features.

GLCP texture features				
	96 to 8	96 to 16	96 to 32	96 to 64
Synthetic	9.33	5.45	2.93	1.55
Brodatz	6.85	4.22	2.33	1.24
Sea ice	2.51	2.04	1.34	0.99
GMRF model parameters				
	96 to 8	96 to 16	96 to 32	96 to 64
Synthetic	28.32	8.50	3.95	1.66
Brodatz	16.44	4.85	2.23	1.24
Sea ice	21.38	7.35	3.19	1.75

Results: Table 2 reports BEB and J for each texture pair. When $n = 8$ and $n = 16$, all of the *GLCP* pairs have a lower BEB as well as a higher J compared to *GMRF* (except for sea ice for $n = 16$). In contrast, for $n = 32$, the *GMRF* has lower BEB and higher J. Separability is relatively stronger given smaller n for *GLCP* features compared to *GMRF* features. However, for large windows, *GMRF* features are more separable relative to *GLCP* features. As a result, if one requires small windows, the *GLCP* method is advocated.

Research Question Three *What is the effect on the estimated features if a window contains multiple textures?*

Theory: For segmentation, some local windows will contain multiple textures.

Method: A reasonable hypothesis is that features derived from a multi-texture window are based on a linear weighting proportional to the spatial extent of each texture. For example, given textures A and B, then $F = a \times F^A + b \times F^B$, where F is the observed texture feature, a and b are the ratios of textures A and B in the window n, and $a + b = 1$.

Results: Three bipartite texture images with vertical center boundaries (synthetic, Brodatz, SAR sea ice) are used for testing. Texture estimates for each image are estimated based on two window sizes ($n = 16$, $n = 32$) for each pixel across fifty randomly selected rows. Selected results are only presented for the sea ice image, $n = 16$, and *GLCP*, given that results for other images

Table 2: Bhattacharyya error bounds (BEB) and Fisher criteria (J) for the indicated texture pairs.

8×8 window size				
	GMRF		GLCP	
	BEB	J	BEB	J
synthetic	3.5×10^{-1}	0.24	2.9×10^{-1}	1.68
Brodatz	3.7×10^{-1}	0.24	2.0×10^{-1}	1.52
sea ice	2.5×10^{-1}	0.55	4.8×10^{-2}	6.58

16×16 window size				
	GMRF		GLCP	
	BEB	J	BEB	J
synthetic	1.7×10^{-1}	4.19	1.3×10^{-1}	4.98
Brodatz	1.3×10^{-1}	2.21	4.4×10^{-2}	5.64
sea ice	7.9×10^{-13}	101.03	1.1×10^{-3}	21.07

32×32 window size				
	GMRF		GLCP	
	BEB	J	BEB	J
synthetic	4.0×10^{-4}	26.33	3.0×10^{-3}	19.75
Brodatz	3.2×10^{-3}	21.85	3.5×10^{-3}	21.54
sea ice	1.5×10^{-51}	462.21	6.7×10^{-8}	58.74

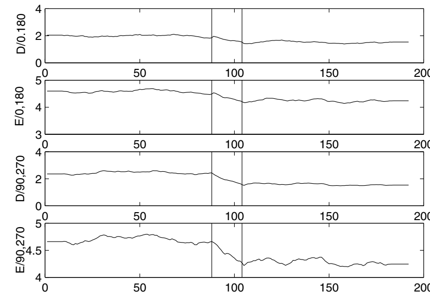

Figure 1: *GLCP* features averaged over 50 arbitrarily selected rows from a SAR sea ice image with two textures separated by a vertical boundary ($n = 16$).

and *GMRF* textures are similar. The vertical lines mark the window centered on the texture boundary. The hypothesis is supported since features change in a linear manner from one texture to another.

Fig. 2 shows the segmentation results of a bipartite image (containing Brodatz paper and pigskin) with an 11 period sinusoidal texture boundary. For both straight and four sinusoidal boundaries given both $n = 8$ and $n = 16$, *GLCP* and *GMRF* and their associated segmentation schemes successfully segment the images (not shown). The segmentation results for both *GLCP* and *GMRF* using a 16×16 window are unsuccessful (Fig. 2(c) and (d)). Fig. 2(e) and (f) are the segmentation results using $n = 8$. In this case, the *GLCP* method produces a much better segmentation than the *GMRF* method. Also, $n = 8$ generates a much better result than $n = 16$ given the *GLCP* texture features. For complex boundaries, both methods may have damaged features estimates which can erode the quality of the segmentation. To minimize the effect of windows containing multiple textures, smaller windows should be used. In such cases, the *GLCP* method should be employed, as supported by the first two research questions.

Fig. 3(a) contains four textures separated by horizontal and vertical sinusoidal boundaries. The larger window size ($n = 32$) allows the *GMRF* method to produce a reasonable segmentation (Fig. 3(f)) compared to a window size of $n = 16$ (Fig. 3(d)). The GLCPs are not able to properly identify the boundary region between the textures (Figs. 3(c) and (e)).

From the segmentation results using different window sizes, one can see that both methods prefer a larger window size to obtain a robust estimation. But, the larger window size may cause a segmentation problem using the *GLCP* method in the boundary area, i.e., the true boundary between textures may be blurred, and sometimes, the pixels along the boundary areas could be distinguished as another texture class. To minimize this boundary problem, the window size should be as small as possible for the *GLCP* method. Using the *GMRF* method, a small window size may ruin the texture model estimation ability. As a result, the window size should be as large as possible for the *GMRF* method. But given complex texture boundaries, a large window size could also damage the estimated *GMRF* models based on the results of the third research question.

4 Summary

There exists a lack of published research comparing unsupervised texture segmentation methods. The goal of this research was to develop a better understanding of the ability of two popular methods for unsupervised image segmentation by considering the role of window size. A number of research questions were posed, producing the following results. *GMRFs* require larger window sizes relative to *GLCPs* to produce stable tex-

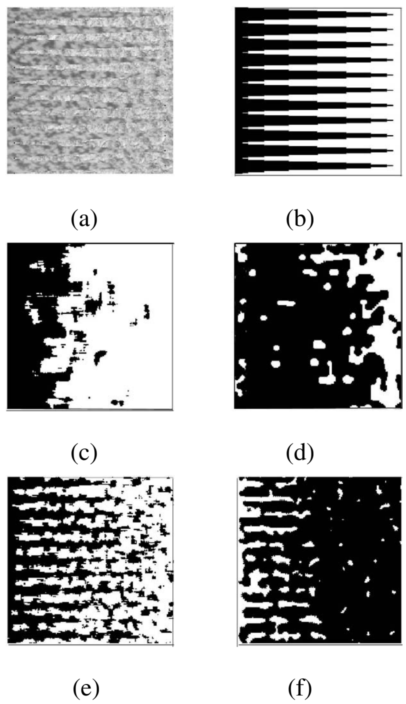

Figure 2: Segmentation of Brodatz mosaic. (a) Original image (b) True segmentation (c) *GLCP* result ($n = 16$). (d) *GMRF* result ($n = 16$). (e) *GLCP* result ($n = 8$). (f) *GMRF* result ($n = 8$).

ture estimates. *GLCPs* produce more separable features for smaller windows relative to the *GMRF*. A window size of 32 was deemed sufficiently large to obtain separable, consistent texture features for the tested textures. However, such a large window can lead to segmentation error due to the higher risk of multiple classes appearing in the same window. Given a window with multiple textures, a region-based weighting of the texture features generates the overall feature. Such a weighting can lead to erroneous boundary estimates and can even identify the boundary itself as belonging to a separate class. The segmentation of classes separated by irregular boundaries will be strongly affected by this process. The texture literature often utilizes convenient texture boundaries, yet complex boundaries coupled with varying local class spatial extents, pose greater challenges in the applied use of segmentation algorithms.

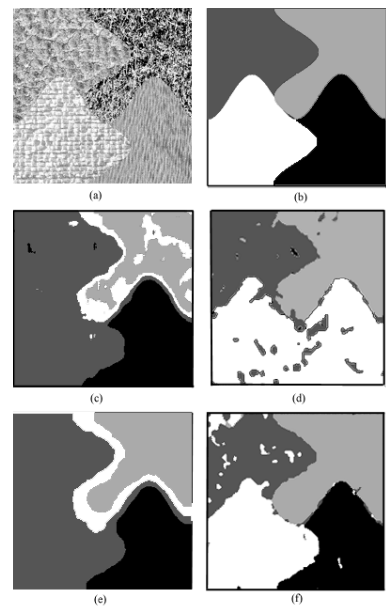

Figure 3: Segmentation of Brodatz texture image. (a) Original image. (b) True segmentation. (c) *GLCP* result ($n = 16$). (d) *GMRF* result ($n = 16$). (e) *GLCP* result ($n = 32$). (f) *GMRF* result ($n = 32$).

References

[1] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society B*, 48:259–302, 1986.

[2] R. Chellappa and S. Chatterjee. Classification of texture using Gaussian Markov random fields. *IEEE Transactions on Acoustics,Speech, and Signal Processing*, 33(4):959–963, 1985.

[3] D. Clausi. Comparison and fusion of co-occurrence, Gabor and MRF texture features for classification of SAR sea-ice imagery. *Atmosphere-Ocean*, 39(3):183–194, 2001.

[4] D. Clausi. An analysis of co-occurrence texture statistics as a function of grey level quantization. *Canadian Journal of Remote Sensing*, 28(1):45–62, 2002.

[5] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley-Interscience, 2001.

[6] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for images classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(6):610–621, 1973.

[7] M. Tuceryan and A. Jain. Texture analysis. In C. Chen, L. Pau, and P. Wang, editors, *Handbook of Pattern Recognition and Computer Vision*, chapter 2.1, pages 235–276. World Scientific Publishing Company, 1993.

Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04)
1051-4651/04 $ 20.00 IEEE

IEEE COMPUTER SOCIETY