MVA'94 IAPR Workshop on Machine Vision Applications Dec. 13-15, 1994, Kawasaki

TEXTURE FEATURES FROM GRAY LEVEL GAP LENGTH MATRIX

Wang Xinli[1,2], Fritz Albregtsen[1], and Bent Foyn[2]

[1] Image Processing Laboratory, Department of Informatics, University of Oslo, P.O.Box 1080, Blindern, N-0316 Oslo, Norway

[2] The Norwegian Pulp and Paper Research Institute, Forskningsveien 3, P.O.Box 24, Blindern, 0313 Oslo, Norway
Email: xinliw@ifi.uio.no or fritz@ifi.uio.no

Abstract

Several texture features are introduced from a proposed higher-order statistical matrix, the gray level gap length matrix (GLGLM). The GLGLM measures the gray level variations in an image. It complements the gray level run length matrix (GLRLM) and is more superior when the number of gray level is large. Features extracted from the weighted GLGLM can be used to estimate the size distribution of the subpatterns. It works much faster than the commonly used κ statistic based on gray level cooccurrence matrices, and provides additional quasi-periodicities. The GLGLM and its features seem very useful in texture segmentation and classification.

Introduction

Texture features are important in pattern recognition, image analysis and synthesis. Higher order statistics describe the relationship of pixels in a more wide extent. They mix the global and local information together into one data set. Therefore, they are more suitable for finding structural characteristics in an image.

We present a set of texture features extracted from a new higher order statistical matrix, the gray level gap length matrix. This matrix reflects directly the size information of texture elements, and the calculation of the matrix and its features are both simple and fast.

Gray Level Gap Length Matrix

The gray level gap length (**GLGL**) method is based on measuring the distribution of gray level gap lengths for each gray level in an image. A gap for gray level g occurs when g is only found at the beginning and the end of a set of consecutive, colinear pixels, while all pixel values in between are either above or below g. The gap length is the distance between these two pixels minus one. Thus, two neighboring pixels with identical gray level have zero gap length. In the case where no pixel with gray level g is found along the searching direction, the gap length is considered as infinite and is omitted.

The gray level gap length matrix (**GLGLM**) is a 2-D array, A(g, l | θ), where g is the gray level, l is the gap length, and θ is the given direction. Given an image of $M \times N$ pixels with G gray levels from 0 to G-1, let $f(i, j)$ be the gray level function at pixel (i, j), and L be the maximum gap length. The element of **GLGLM** at angle θ is defined as:

$$A(g, l|\theta) = \#\{ (i,j) \mid f(i,j) = f(i+x, j+y) = g, \quad f(i+u, j+v) \neq g, \quad u<x \;\&\; v<y \} \tag{1}$$

where "#" stands for "the number of", and

$$x = (l+1)\cos\theta, \qquad y = (l+1)\sin\theta,$$

$$0 \le g \le G-1, \qquad 0 \le l \le L, \qquad 0^o \le \theta \le 180^o$$

An average matrix from **GLGLM**s obtained in different directions will give a rotation invariant representation of the texture patterns.

Texture Features Extracted from GLGLM

As Mentioned in [5], the GLGLM has a relation to the gray level cooccurrence matrix (**GLCM**), but it is not a variant of GLCM. The diagonal elements of cooccurrence matrices are the counting of the repeated pixel values at a given displacement d. It can not, however, distinguish whether the distance d really represents a gap, or it is just a combination of several small gaps. This information is lost once the statistic measurements finish.

On the other hand, the GLGL method can be viewed as a complement to the Gray Level Run Length (**GLRL**) method. A gray level run length is the length of some collinearly adjacent pixels having the same gray value [3].

From **Figure 1** we see that the run length and the

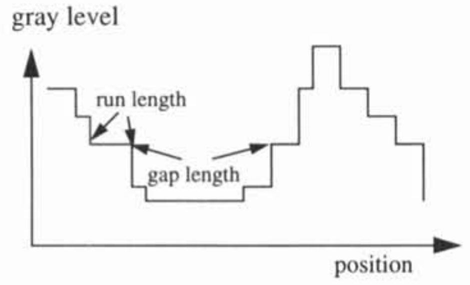

Figure 1
The relationship between run length and gap length.

gap length are different measurements of the image, although both methods are approaches of higher-order statistics. The GLRLM measures plateaus of the image whereas the GLGLM reflects peaks and valleys of the image. In fact, the gap length matrix obtains more information than the run length matrix. The GLRLM will tend to have many small runs when the number of gray levels increases. The structure information is therefore lost. On the contrary, the GLGLM will still catch the variations in the image. So, it can reflect the texture information to a more deep extent.

When the image is re-quantified into two levels, both methods catch the similar information, except that the gap statistics for the white pixels is the runs for the black pixels and vice versa. An important advantage of the gap length method, however, is that it ignores the information of the background.

Figure 2a shows a 128×128 pixel binary image of a ring, having a thickness of 3 pixels and an inner diameter of 74. Figure 2b is the gap length statistics along the horizontal axis. Here the solid line and dashed line are the gap length statistics for black pixels and white pixels respectively. The highest peak for black pixels is at gap length 74, which is the inner diameter of the ring, while the most frequent gap length for white pixels is 3. Figure 2c is the run length statistics. Although it measures the same image along the same direction, it only catches the correct width of the ring (the highest point in the dashed line). The most frequent run length for white pixels (the highest peak in the solid line) is 128, which is the width of the image. Even the second highest peak has nothing to do with the ring itself. It is just about the shortest distance between the ring and the boundaries of the image. Thus, we see that gap length statistics are more relevant for local structure detection, as it is less influenced by the background. This is specially useful when the windows used for texture analysis are relatively small.

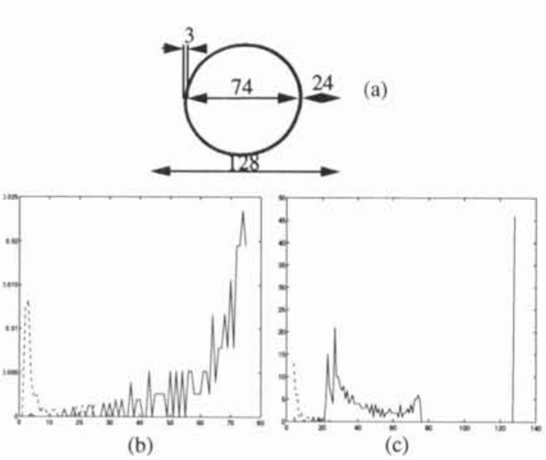

Figure 2
(a) A binary image of a ring.
(b) Its gap length statistics (dashed line for white level and solid line for black level).
(d) Its run length statistics (dashed line for black level and solid line for white level).

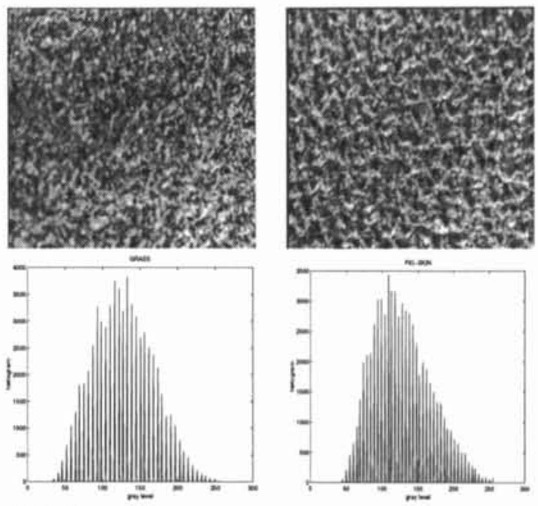

Figure 3
Top: Two Brodatz images, "Grass" and "Pig-skin".
Bottom: Their histograms with mean of 128 and standard deviation of 40.

Several texture features can be extracted from the GLGLM.

Assume that

-- $A(g, l \mid \theta)$ is the (g, l)th element of the gap length matrix for a direction θ,

-- G is the number of gray levels,

-- L is the longest gap length,

-- n is the number of pixels in the image.

Short Gaps Emphasis:

$$SGE = \sum_{g=0}^{G-1}\left(\sum_{l=0}^{L}\frac{A(g, l|\theta)}{l^2}\right) \Big/ \sum_{g=0}^{G-1}\left(\sum_{l=0}^{L} A(g, l|\theta)\right) \quad (2)$$

By dividing each gap length value by the square of its length, short gap lengths are emphasised. The denominator is the total number of gaps in the image.

Long Gaps Emphasis:

$$LGE = \sum_{g=0}^{G-1}\left(\sum_{l=0}^{L} l^2 A(g, l|\theta)\right) \Big/ \sum_{g=0}^{G-1}\left(\sum_{l=0}^{L} A(g, l|\theta)\right) \quad (3)$$

Here we multiply each gap length value by the square of its length, in order to give higher weight to the long gaps.

Gray Level Fluctuation:

$$GLF = \sum_{g=0}^{G}\left(\sum_{l=0}^{L} A(g, l|\theta)\right)^2 / \sum_{g=0}^{G-1}\left(\sum_{l=0}^{L} A(g, l|\theta)\right) \quad (4)$$

High gap length value will contribute most to this feature. The GLF feature will have its lowest value if the number of gaps are evenly distributed over all gray levels.

Gap Length Non-uniformity:

$$GLN = \sum_{l=0}^{L}\left(\sum_{g=0}^{G-1} A(g, l|\theta)\right)^2 / \sum_{g=0}^{G-1}\left(\sum_{l=0}^{L} A(g, l|\theta)\right) \quad (5)$$

The GLN feature will have its lowest value if the number of gaps are evenly distributed over all gap length.

Gap Percentage:

$$GP = \frac{1}{n}\sum_{g=0}^{G-1}\left(\sum_{l=0}^{L} A(g, l|\theta)\right) \quad (6)$$

This feature is the ratio between the total number of observed gaps in the image and the total number of possible gaps if all gap had a length of one.

Low Gray level Gaps Emphasis:

$$LGGE = \sum_{g=0}^{G-1}\left(\sum_{l=0}^{L} \frac{A(g, l|\theta)}{g^2}\right) / \sum_{g=0}^{G-1}\left(\sum_{l=0}^{L} A(g, l|\theta)\right) \quad (7)$$

High Gray level Gaps Emphasis:

$$HGGE = \sum_{g=0}^{G-1}\left(\sum_{l=0}^{L} g^2 A(g, l|\theta)\right) / \sum_{g=0}^{G-1}\left(\sum_{l=0}^{L} A(g, l|\theta)\right) \quad (8)$$

LGGE and HGGE are introduced in order to distinguish textures that are similar according to their SGE and LGE features, but differ in gray level distribution of the gaps.

Table 1

Requantized gray levels		4		16	
Name of Image		Grass	Pig-skin	Grass	Pig-skin
Fea-tures	SGE	0.796	0.787	0.383	0.372
	LGE	65.9	73.7	272	313
	GLF	27350	27578	7182	7329
	GLN	39582	38608	9081	8534
	GP	0.985	0.984	0.954	0.952
	LGGE	0.207	0.193	0.0185	0.0181
	HGGE	6.617	6.817	77.86	77.56

These features seems analogical to those of the GLRLM [2], but they have different meanings and measure different aspects of the image. Figure 3 shows two Brodatz images [1], "Grass" and "Pig-skin", with the histograms linearly rescaled at mean 128 and standard deviation 40. **Table 1** shows some preliminary results obtained from these two images. Although the first order statistics are same, such that the differences between GP, LGGE, HGGE are small, one can still discriminate the two textures by the other features.

We can further expect that a combining feature set from GLGLM and other statistic methods may result in an improvement for texture classification and analysis.

Finding Pattern Sizes by GLGLM

Finding the size distribution of the subpatterns is important for texture analysis and synthesis. Since the gap length method measures the waves in an image, we can extract the size information from the GLGLM.

As in the case of **GLCM** [4], the number of occurrences decreases as the gap length increases. For statistics at each gap length to be comparable, a weighting has to be done.

The weighted GLGLM, $A_w(g, l \mid \theta)$, can be obtained by:

$$A_w(g, l|\theta) = \frac{A(g, l|\theta)}{W(l)} \quad (9)$$

where the weighting function $W(l)$ varies according to gap length l. For the analysis along the x or y direction only, the weighting function can be chosen as follows:

When $l \le G$

$$W(l) = (M-x) \cdot (N-y) \quad (10)$$

When $l > G$

$$W(l) = W_1(x) \cdot W_2(y) \quad (11)$$

where

$$W_1(x) = \begin{cases} ((M/x) - 1) \cdot (G-1) & x \ne 0 \\ M & x = 0 \end{cases} \quad (12)$$

$$W_2(y) = \begin{cases} ((N/y) - 1) \cdot (G-1) & y \ne 0 \\ N & y = 0 \end{cases} \quad (13)$$

The special consideration for the case where $l > G$ is due to the fact that no repetition of gray level g is allowed in a gap length l. So, the weighting function is approximately the maximum possible number of gaps of a length l in an image.

Normally, only a few gray levels are used in the weighted GLGL matrix, and a shading correction is often useful to increase the possibility for finding significant gaps. Histogram linear extension is performed first in order to extend the dynamic range of the gray levels. Then the image is re-quantified into 2 to 8 gray levels. Since adjacent gray levels normally catch similar structural information, a sum of gap lengths among adjacent gray levels will help to find the size distribution of the structure.

For example, the size distributions for darker gray levels and brighter gray levels can be obtained by:

$$S_{black}(l) = \sum_{g=0}^{\mu_0 - 1} A_w(g, l) \qquad (14)$$

$$S_{white}(l) = \sum_{g=\mu_0}^{G-1} A_w(g, l) \qquad (15)$$

$$\mu_0 = \frac{\sum_{g=0}^{G-1} h(g) \cdot g}{\sum_{g=0}^{G-1} h(g)} \qquad (16)$$

where $h(g)$ is the histogram value of gray level g, G is the number of gray levels and $0 \le l \le L$.

For the texture of solid subpatterns, this method can be used to estimate the quasi-periodicity. Figure 4a shows a frequently quoted natural image, 'loose burlap' [1]. The histogram extension was performed first, then the image was re-quantified into $G = 8$ gray levels. The detail of the texture cells is shown in Figure 4b.

The weighted GLGLM is calculated along the horizontal direction, then it is summed into two levels as in (14) and (15). The results are shown in Figure 4c. The highest peaks for black and white gray levels indicate the estimated quasi-periodicities. The period is just the sum of these two quasi-periodicities.

A commonly used κ statistic is also applied as a comparison. Thirty gray level cooccurrence matrices from displacement $d = 1$ to 30 are calculated along the same direction. The extracted distribution of κ for each d is shown in Figure 4d. The period is the displacement where the highest peak occurs.

Figure 4 shows us that the gap length statistic catch the same period value as the κ does, 5 pixels, while the former also gives the correct estimation for size of the black pixels, 3, and the white pixels, 2. The ratio of CPU time in this case is about 1/30 for the gap length statistic to κ.

This property of the weighted GLGLM features can also be used to measure the size distribution of grains, or to decide the window size for texture segmentation [5]. In the former case, the estimated grain sizes can be used directly in granulometry for deciding the size of structuring elements, where the time and memory saving is tremendous. The gap length statistics may also serve as a signature of the image.

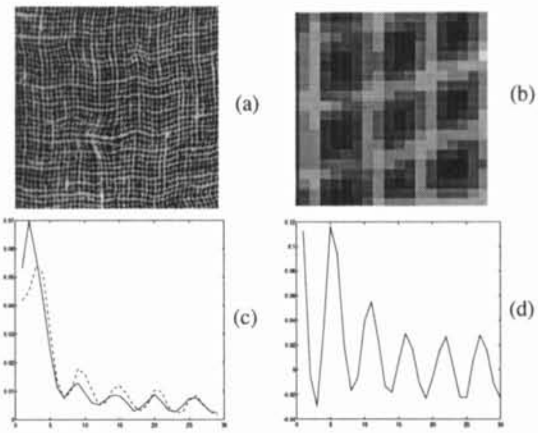

Figure 4

(a) Original 'burlap' image.

(b) A part of the 3-bit representation of 'burlap'.

(c) The gap length statistics of (a), where the solid line is the gap length of black gray levels and the dashed line is that for white gray levels. The horizontal axis is the gap length.

(d) The κ statistic as the vertical axis and the displacement as the horizontal axis for (a).

Conclusion

The proposed gray level gap length method extracts texture statistic information in a new and efficient way. It catches more information than the gray level run length method. It works much faster than the commonly used κ statistics and provides additional quasi-periodical information. The texture features extracted from it seem quite valuable for texture segmentation and classification. The GLGLM indeed opens a wide way for texture analysis.

References

[1] Brodatz P.: "*Textures: A Photographic Album for Artists and Designers*", Reinhold, New York, 1986.

[2] Chu A. et al. "*Use of Grey Value Distribution of Run Lengths for Texture Analysis*" Pattern Recognition Letters 11, pp. 415-420, June 1990.

[3] Galloway M. M.: "*Texture Analysis Using Gray Level Run Lengths*", Comp. Graphics and Image Processing 4, pp. 172-179, 1975.

[4] Parkkinen J. and k. Selkåinaho: "*Detecting Texture Periodicity from the Cooccurrence Matrix*", Pattern Recog. Letters, 11, pp.43-50, 1990.

[5] Wang X., F. Albregtsen and B. Foyn: *Grey Level Gap Length Matrix: A New Approach for Texture Analysis*, to appear in Proceeding of 3rd ICARV'94, Singapore, November 1994.