Spatial Statistics for Content Based Image Retrieval

Suryani Lim, Guojun Lu
Gippsland School of Computing and Information Technology
Monash University, Australia
Suryani.Lim,Guojun.Lu@infotech.monash.edu.au

Abstract

A Content Based Image Retrieval (CBIR) system retrieves relevant images from an image database. Over the years, several methods have been proposed to extract these features. Previous researches show that the effectiveness of a CBIR system increases when spatial relationship of colours is considered. In this paper we propose using the Looseness parameter from geostat, a branch of statistics which deals with geographical data, to describe the global spatial relationship of colours. Spatial Chromatic Histogram (SCH) is another method which also measures the global spatial relationship of colours. However, the spatial measurement of SCH is size variant, the spatial measurement of geostat is size invariant. We will analyse and compare the performance of geostat and SCH.

1. Introduction

Computer vision is a challenging task. While it is easy for human to recognise an object, it is extremely difficult to "teach" computers to recognise an object. In the early nineties Swain and Ballard developed a very elegant method which identifies an object based on the colour distribution of the object using colour histogram [7]. The method they developed require massive reduction of over 2 million colours to several hundred bins of colours, a process known as colour quantisation. Each bin contains colours which are perceptually similar. The colour histogram is the count of pixels of the quantised colours. The similarity of two objects are measured using their colour histograms. In other words, an object is indexed using its colour histogram. The solution has also been extended from object recognition to image recognition so that we can retrieve images based on the content of the images not on the text annotation of the images.

In this paper we propose using the Looseness parameter from geostat, a branch of statistics which deals with geographical data, to describe the global spatial relationship of colours. Spatial Chromatic Histogram (SCH) is another method which also measures the global spatial relationship of colours. We will analyse and compare the performance of geostat and SCH. Section 2 describes previous work on CBIR based on colours which incorporate spatial relationship of the quantised colours (to avoid repetition the term colours are used to represent the term quantised colours). SCH and the effect of size variant spatial measurement of SCH will also be described in the same section. Section 3 describes the Looseness parameter from geostat and provides analysis on how it can solve the problem of SCH. Section 4 describes the experiment set up and Section 5 discusses the results. Section 6 describes future work and Section 7 concludes the paper.

2. Previous Work

The main problem with colour histogram indexing is that it does not take the spatial relationship of the colours into consideration. Images in Figure 1b-d are visually different but their indexes are the same. Many attempts have been made to rectify this problem by using indexes which incorporate spatial relationship of the colours. We can categorise the level of spatial relationship into two categories: local and global. Indexes which consider the spatial relationship of the colours in a certain area of the image is considered as having local spatial relationship. Indexes which consider the spatial relationship of the colours of the whole image is considered as having global spatial relationship. In next sections, we review two indexing techniques having global spatial relationship: Annular Colour Histogram and Spatial-Chromatic Histogram.

2.1. Annular Colour Histogram

Rao et al [6] introduced annular colour histogram. Let A_i be the count of pixels in bin i. Let $C_i = (x_i, y_i)$ be the centroid of bin i, x_i and y_i are defined as:

$$x_i = \frac{1}{A_i} \sum_{(x,y)\in A_i} x; \;\; y_i = \frac{1}{A_i} \sum_{(x,y)\in A_i} y \qquad (1)$$

Proceedings of the International Conference on Information Technology: Computers and Communications (ITCC'03)
0-7695-1916-4/03 $17.00 © 2003 IEEE

Query Image

		colour	hist.	σ	L
(a)		red	1.00	104.51	1.05
		yellow	0.00	0.00	0.00

Other Images

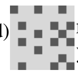
		colour	hist.	σ	L	dis(SCH)	dis(Geostat)
(b)		red	0.25	52.25	1.05	4.45	2.86
		yellow	0.75	116.85	1.75		
(c)		red	0.25	63.08	1.52	3.99	3.68
		yellow	0.75	114.99	1.60		
(d)		red	0.25	107.46	4.43	2.91	6.28
		yellow	0.75	102.65	1.35		

Figure 1. Traditional colour histogram does not take into account the spatial distribution of colours. Image b-d have exactly the same colour histogram although visually they look very different. SCH (σ_{red}) and geostat (L_{red}) take into account the spatial distribution of the colours so the values are sensitive to the spatial distribution of the colours.

Let r_i be the radius of bin i:

$$r_i = max_{(x,y)\in A_i} \sqrt{(x-x_i)^2 + (y-y_i)^2} \qquad (2)$$

Divide each radius r_i into N and draw N concentric circles centred at C_i. A_{ij} is the count of pixels having colour bin i inside circle j, it is the annular colour histogram. Figure 2 illustrates the indexing of annular colour histogram of black pixels of an image where N = 4. The dissimilarity metric is the Euclidian distance of the annular colour histogram. The size of the index is $N\times$ the size of a traditional histogram.

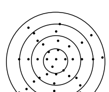


Figure 2. Annular colour histogram

2.2. Spatial-Chromatic Histogram

Cinque et al [3] introduced spatial-chromatic histogram (SCH) to describe how pixels of identical colour are distributed in the image. Let $C_i = (x_i, y_i)$ be the centroid of bin i (x_i and y_i are calculated using Formula 1). Let σ_i be the standard deviation of bin i from C_i and it is defined as:

$$\sigma_i = \sqrt{\frac{\sum_{p\in A_i} d(p, C_i)}{A_i}^2} \qquad (3)$$

$$d(p, C_i) = \sqrt{(x-x_i)^2 + (y-y_i)^2} \qquad (4)$$

where p is the position of a pixel at (x, y). σ_i measures the square root of the average squared distance of pixels in bin i from its centroid.

The SCH index for image Q is $(h_1^Q, \sigma_1^Q, \ldots, h_M^Q, \sigma_M^Q)$, where h_1^Q is the histogram of bin 1 of image Q, σ_1^Q is the spatial distribution of bin 1 of image Q and M is the number of bins.

The similarity metric of image Q and I is defined as:

$$f_s(Q, I) = \sum_{i=1}^{M} min(h_i^Q, h_i^I) \times \frac{min(\sigma_i^Q, \sigma_i^I)}{max(\sigma_i^Q, \sigma_i^I)} \qquad (5)$$

The similarity metric is made up of two parts. The first part, $min(h_i^Q, h_i^I)$, is histogram intersection and it measures the similarity between h_i^Q and h_i^I. The second part ($\frac{min(\sigma_i^Q,\sigma_i^I)}{max(\sigma_i^Q,\sigma_i^I)}$) measures the similarity of the spatial distribution of the pixels of a bin. Since σ describes how disperse the pixels in the bin are, $\frac{min(\sigma_i^Q,\sigma_i^I)}{max(\sigma_i^Q,\sigma_i^I)}$ is more meaningful than $min(\sigma_i^Q, \sigma_i^I)$ or $abs(\sigma_i^Q - \sigma_i^I)$. If σ_i^Q and σ_i^I are identical the score is 1. If σ_i^Q and σ_i^I are different the score deviates from 1 towards 0. So indexes of two images are similar only if their colour histograms and spatial distributions are similar. Please note that f_s described in Formula 5 is modified from the f_s described in [3]. The f_s described in [3] is rotation **variant** while f_s described in Formula 5 is rotation **invariant**.

The dissimilarity metric is simply:

$$f_d(Q, I) = \frac{1}{0.1 + f_s(Q, I)} \qquad (6)$$

Figure 1 shows the dissimilarity scores of a query image with three other images. Note that the spatial distribution for each colour is different in each image and so are the dissimilarity scores.

SCH is more efficient than annular colour histogram because the size of an SCH index is smaller than the size of an annular colour histogram index. However, there is a problem with σ. Because σ is the square root of average squared distance from its centroid, a large σ could mean two things: many pixels with compact structure or a few pixels with loose structure. In other words, it is size variant. This problem can be illustrated in Figure 1. Visually Figure 1a is closest to Figure 1b but $\sigma_{red}^{Figure\ 1a}$ is furthest away from $\sigma_{red}^{Figure\ 1b}$. The retrieval results using Figure 1a as the query image can be seen in Figure 3. The image which is visually closest to the query image is ranked last.

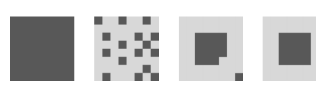


Figure 3. Retrieval results of SCH using left image as a query. The image which is visually closest to the query image is ranked last.

3. Using Geostat for Image Retrieval

Geographical statistics (geostat) is a branch of statistics for studying the "distribution of populations over territories" [2]. One parameter in geostat called Looseness is particularly useful for comparing the spatial distribution of population under study.

3.1. Basic Terminologies

In geostat, the feature we are measuring is known as a geoset. The map of a geoset is made up of a number of regions and each region is identified by its spatial location and number of elements. A geoset can be formally defined as:

$$^{g}H = \{x_{_{H}s}, y_{_{H}s}, w_{_{H}s}\}$$

where ^{g}H is the geoset H, $x_{_{H}s}$ and $y_{_{H}s}$ are the x and y coordinates of subregion s and $w_{_{H}s}$ is the number of element in subregion s.

Geo-means

Figure 4 shows a map of 12 elements. To calculate the geo-means, divide the map into subregions then label the centroid of each subregion with coordinates. The geo-means $(\bar{x}, \bar{y})$ is defined as:

$$\bar{x} = \frac{1}{w_t}\sum_{s=1}^{S} w_s\, x_s;\ \bar{y} = \frac{1}{w_t}\sum_{s=1}^{S} w_s\, y_s \tag{7}$$

where S is the number of subregions, x_s and y_s are the x and y coordinates of subregion s, w_s is the number of elements in subregion s and w_t is the total number of elements.

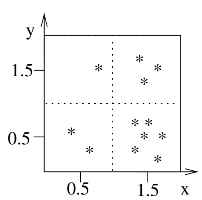


Figure 4. A map divided into 4 subregions. The location of each element is marked with with an '*'.

Looseness

To calculate looseness (L), we need to calculate the distance variance (V) of the geoset. Distance variance is a measure of the variance of distance of each subregion from the geo-means and is defined as:

$$V = \frac{1}{w_t}\sum_{s=1}^{S} w_s\, d((x_s, y_s), \text{geo-means})^2 \tag{8}$$

$$L = \frac{1}{Area} 2\pi V \tag{9}$$

where $d((x_s, y_s), \text{geo-means})$ is defined in Formula 4, $Area$ is the area of the net territory over which population is distributed. In Figure 4, the elements are distributed over 4 subregions and the size of each subregion is one square unit, therefore $Area$ is 4.

3.2. Using Looseness to Measure Spatial Relationship of Colours

To use L for CBIR, the geosets are the quantised colours and we are interested in the spatial distribution of pixels having identical colour (quantised into the same bin). If an image has $row \times col$ pixels and we divide the image into $row \times col$ subsections, then w_{is} is always 1. As a result $Area_i$ is the number of pixels in bin i and geo-means$_i$ is C_i.

The definition of V_i (Formula 8) and L_i (Formula 9) can now be simplified to:

$$V_i = \frac{1}{A_i}\sum_{p \in A_i} d(p, C_i)^2 \tag{10}$$

$$L_i = \frac{1}{A_i} 2\pi V_i \tag{11}$$

where C_i is the centroid of bin i at (x_i, y_i) and it is calculated using Formula 2, $d(p, C_i)$ is defined in Formula 4, p is the position of a pixel at (x, y) and A_i is the number of pixels in bin i. The geostat index for image Q is $(h_1^Q, L_1^Q, \ldots, h_M^Q, L_M^Q)$, where h_1^Q is the histogram of bin 1 of image Q, L_1^Q is the spatial distribution of bin 1 of image Q and M is the number of bins.

3.3. Interpretation of Looseness

There are two ways we can interpret L. With reference to geostat, V measures the dispersion, therefore dividing L by the net territory measures the dispersion per unit area.

With reference to SCH, L can also be defined in terms of σ:

$$L_i = \frac{1}{A_i} 2\pi\sigma_i^2 \tag{12}$$

Recall that σ is the square root of the average squared distance of pixels from the centroid and its value could be similar for bins having different number of pixels. Dividing σ^2 by the number of pixels in the bin normalises the dispersion to the number of pixels in the bin. The main difference between L and σ is L is size invariant but σ is not.

It can be seen from Figure 1 that as L_{red} approaches 1, red is compact and as it departs from 1 it is less compact. The most compact structure is that of a circle.

Since both σ and L describe the spatial distribution of a colour, the similarity metric defined in Formula 5 and dissimilarity metrics defined in Formula 6 are also used to measure the similarity and dissimilarity of two indexes but σ is replaced with L.

0-7695-1916-4/03 $17.00 © 2003 IEEE

The retrieval results using geostat can be seen in Figure 5. The left most image is the query image and the image visually closest to the query image is ranked first.

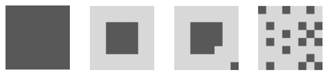


Figure 5. Retrieval results of geostat using left image as a query. The image which is visually closest to the query image is ranked first.

4. Experiment Setup

We conducted the experiments using GNU Image Finding Tool (GIFT) [5]. GIFT is an open source client server CBIR framework and it can be downloaded from the URL given in[1]. GIFT has some indexing algorithms but it also allows users to specify their own indexing algorithms by writing plug-in.

We indexed over 10,000 images using SCH and geostat in HSV colour space. The colour space is uniformly quantised into 21 levels of hue, 3 levels of saturation and value giving a total of 189 bins. The metric for evaluating dissimilarity between two images is given in Formula 6.

Nine images from the database are selected as query images. For each query image, we established a set of ground truth and use precision(P) and recall(R) graph to evaluate the effectiveness: $P = \frac{r}{N}$ and $R = \frac{r}{total\ relevant}$ where r is the number of relevant images retrieved, N is the number of images retrieved, $total\ relevant$ is the total number of relevant images. In evaluating the effectiveness of two CBIR systems, the one which gives higher precision value at the same recall value is the more effective system.

5. Results

5.1. Analysis

Figure 6 shows the top ten retrieval results of a query (the top left image is the query image). The retrieval results using geostat is more effective than SCH because the top 8 results in SCH are relevant but the top 10 results in geostat are relevant.

The effect of the problem with σ parameter from SCH described in Section 2.2 can be seen in Figure 6. "Butterfly" which is visually very different from "Sushi" has higher ranking than other sushi images. This happened because σ is size **variant**. For example, Figure 7 shows the pixels of bin number 17 for "Sushi" and "Butterfly". The spatial relationship of the colours are different and the difference is reflected well in the Looseness parameter but not in σ. Consequently, "Butterfly" is ranked higher than other sushi images in SCH.

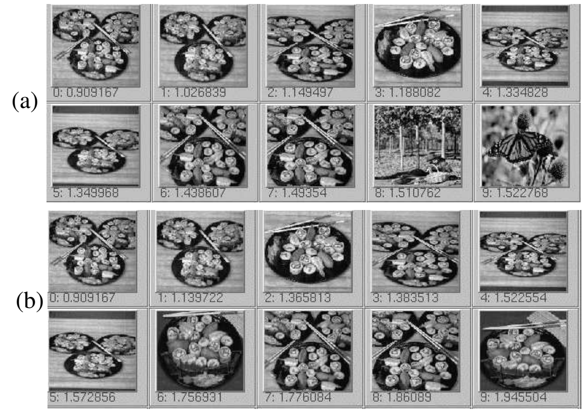


Figure 6. Retrieval results of query image "Sushi" using (a) SCH (b) geostat

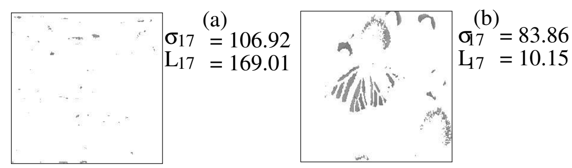


Figure 7. Pixels in bin 17 for (a) "Sushi" and (b) "Butterfly". The spatial relationship of the colours are different and the difference is reflected well in the Looseness parameter but not in σ.

5.2. Colour Quantisation Error

We found a problem when measuring the global spatial relationship of colours. We assume that perceptually similar colour is quantised into the same bin so images of which quantised colours with similar Looseness parameter should be similar. In reality, perceptually similar colours may not be quantised into the same bin. The evidence is clear in query image "Brazil"(see Figure 8a and b). To human, the yellow t-shirt is a compact structure but to the computer after colour quantisation it is not a compact structure. Figure 9 shows the pixels of bin number 35 for "Brazil" and "Brazil 2" ("Brazil 2" is the image ranked 4^{th} in Figure 8a and ranked 9^{th} in Figure 8b). The spatial relationship of bin number 35 in both images are indeed very different and this is reflected well in the Looseness parameter. This problem could be solved by using a coarser colour quantisation but it is also known that coarse colour quantisation reduces the effectiveness of colour based CBIR. This colour quantisation error problem is common to all colour based CBIR.

5.3. Overall Results

Figure 10 is the Recall and Precision graph of the results averaged over nine queries. The effectiveness of geostat is comparable to or slightly better than SCH. In some queries, like "Sushi", geostat performs better than SCH. But in other

Figure 8. Retrieval results of query image "Brazil" using (a) SCH (b) geostat

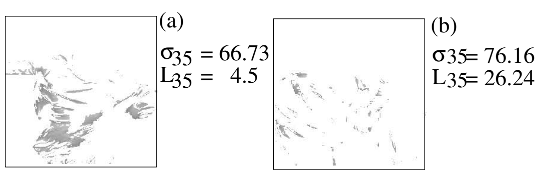


Figure 9. Pixels in bin 35 for (a) "Brazil" and (b) "Brazil 2". The spatial relationship of the colours are different and the difference is reflected well in the Looseness parameter but not in σ.

queries, like "Brazil", SCH performs better because similar σ is obtained on very different "wrongly" quantised colours.

6. Future Work

Geostat measures the spatial relationship of colours globally. The ability to describe spatial relationship of colours seems powerful but it is not clear how it will perform against indexes which describe local spatial relationship such as Colour Autocorrelogram [4] and VQ [8]. We intend to study the performance of geostat against Colour Autocorrelogram and VQ.

7. Conclusion

We introduced the Looseness (L) parameter from geostat as a method of measuring the global spatial relationship of colours in an image. We also analysed and compared the performance of geostat with SCH. The main difference between L from geostat and σ from SCH is that Looseness is size invariant and σ is not. For some query images, geostat performs better than SCH and for some query images, SCH

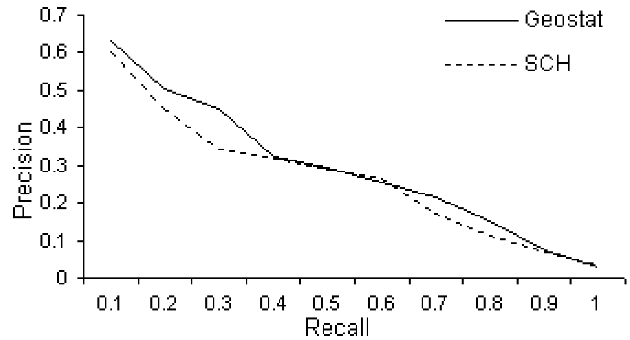


Figure 10. Average Recall and Precision Graph averaged over seven query images

performs better than geostat. Overall, the effectiveness of geostat is comparable to or slightly better than SCH. We have also explained why geostat, which has more accurate spatial measurement than SCH may sometimes have lower effectiveness than SCH.

Acknowledgment

The authors wish to thank Mr Shyh Wei Teng of the Gippsland School of Computing and IT, Monash University for providing the image database and the ground truths.

References

[1] GIFT : The GNU Image-Finding Tool. http://www.gnu.org/software/gift.

[2] R. Bachi. *New Methods of Geostatistical Analysis and Graphical Presentation*. Kluwer Academic Publishers, 1999.

[3] L. Cinque, S. Levialdi, K. Olsen, and A. Pellicano. Color-based image retrieval using spatial-chromatic histograms. In *Multimedia Computing and System, IEEE Intl Conf on*, volume 2, pages 969–973, 1999.

[4] J. Huang, S. Kumar, M. Mitra, W. Zhu, and R. Zabih. Image indexing using color correlograms. In *IEEE CVPR*, Puerto Rico, Jun 1997.

[5] H. Müller, W. Müller, S. Marchand-Maillet, D. Squire, and T. Pun. Automated benchmarking in content-based image retrieval. In *Proc. of the 3rd Intl Workshop on Multimedia Info. Retrieval (in conjunction with ACM MM'01)*, Canada, 2001.

[6] A. Rao, R. Srihari, and Z. Zhang. Spatial color histograms for content-based image retrieval. In *11th IEEE Intl Conf on Tools With AI*, 1999.

[7] M. Swain and D. Ballard. Color indexing. *Intl Journal of Computer Vision*, 7(1):11–32, 1991.

[8] S. Teng and G. Lu. Performance study of image retrieval based on vector quantization. In *Intl Conf on Intelligent Multimedia and Distance Education(ICIMADE)*, pages 76–80, North Dakota, USA, 2001.

0-7695-1916-4/03 $17.00 © 2003 IEEE