SPIRAL BIT-MAP SIGNATURE

FOR

CONTENT-BASED IMAGE RETRIEVAL

WANG HUI HUI

A thesis submitted
in full fulfillment of the requirement for the degree of
Master of Science in Information Technology

FACULTY OF INFORMATION TECHNOLOGY
UNIVERSITI MALAYSIA SARAWAK
2003

UNIVERSITI MALAYSIA SARAWAK

BORANG PENGESAHAN STATUS TESIS

JUDUL: **SPIRAL BIT-MAP SIGNATURE FOR CONTENT-BASED IMAGE RETRIEVAL**

SESI PENGAJIAN : 2001/2002

Saya WANG HUI HUI
(HURUF BESAR)

mengaku membenarkan tesis * ini disimpan di Pusat Khidmat Maklumat Akademik, Universiti Malaysia Sarawak dengan syarat-syarat kegunaan seperti berikut:

1. Tesis adalah hakmilik Universiti Malaysia Sarawak
2. Pusat Khidmat Maklumat Akademik, Universiti Malaysia Sarawak dibenarkan membuat salinan untuk tujuan pangajian sahaja
3. Membuat pendigitan untuk membangunkan Pangkalan Data Kandungan Tempatan
4. Pusat Khidmat Maklumat Akademik, Universiti Malaysia Sarawak dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi
5. ** sila tandakan (√)

☐	SULIT	(mengandungi maklumat yang berdarjah keselamatan atau kepentingan seperti termaktub di dalam AKTA RAHSIA RASMI 1972)
☐	TERHAD	(mengandungi maklumat Terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
√	TIDAK TERHAD	

Disahkan oleh

WANG HUI HUI
(TANDATANGAN PENULIS)

PROF. MADYA NARAYANAN KULATHURAMAIYER
(TANDATANGAN PENYELIA)

Alamat Tetap:
27, Taman Penrissen,
Kota Sentosa,
93250 Kuching, Sarawak

Tarikh: 5 June 2003

Tarikh: 5 June 2003

Catatan : * Tesis dimaksudkan sebagai tesis bagi Ijazah Doktor Falsafah, Sarjana dan Sarjana Muda
** Jika Tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh tesis ini perlu dikelaskan sebagai SULIT atau TERHAD

DECLARATION

No portion of the work in this report has been submitted in support of an application for another degree or qualification of this or any other university or institution of higher learning

Wang Hui Hui
01-02-0638
5 JUNE 2003

ACKNOWLEGEMENT

I would like to take this opportunity to firstly thank to my supervisor, Assoc. Prof. Narayanan Kulathuramaiyer for all the invaluable guidance and encouragement provided throughout this research. I would also like to thank my co-supervisor, Dr. Abdelhamid Abdesselam, for his invaluable technical guidance and ideas. I would also like to thank Mr.Edwin Mit for his advice and support during the initial research period.

I would like to thank all of my friends for their kindness and useful discussions, especially during difficult times.

I would like to acknowledge the support received from the *Pelan Tindakan Pembangunan Teknologi Perindustrian* (PTPTP) and the Graduate Studies & Research Support Division, Universiti Malaysia Sarawak.

Finally, I am deeply grateful to my beloved family for their unconditional support and encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLE

ABSTRACT

Images are generated in many application areas such as image processing, multimedia, digital libraries, remote sensing, astronomy, database applications and other related area at a fast rate. A content-based image retrieval (CBIR) systems needs to provide an effective and efficient access to image databases, based on visual content which includes colour, shape, texture and spatial features.

The research has focused on enhancing low-level image retrieval features in order to address transform invariance. In addition, this research has introduced a promising novel technique, which integrates both color and spatial information together in order to overcome the limitations of current systems.

This research proposed the spiral bit-map signature for representing the spatial feature in colour-spatial image retrieval. The approach has been proven to be rotation invariant as well as able to locate contained image (partial-image) within other images.

ABSTRAK

Penghasilan imej di dalam aplikasi seperti pemprosesan imej, multimedia, perpustakaan berdigit, penderiaan jauh, aplikasi penyimpanan data dan aplikasi lain pada kadar yang mendadak. Satu sistem penolehan inej diperlukan untuk menyediakan cara penolehan imej yang lebih cekap dan berkesan berdasarkan kepada kandungan visual imej termasuk ciri warna, rupabentuk, corak dan makluniat penaburan..

Penyelidikan ini menumpukan perhatian kepada memperkukuhkan ciri peringkat rendah dalam perolehan imej untuk mengatasi transformasi tidak varian. Selain itu, penyelidikan ini telah memperkenalkan satu teknikyang mengabungkan ciri warna dan maklumat penaburan bagi mengatasi batasan yang dialami oleh sistem yang ada pada masa kini.

Kajian ini telah memperkenalkan kaedah pusaran "bit-map'' untuk membentuk perwakilan bagi maklumat penaburan dalam teknologi perolehan imej berdasarkan warna and penaburannya. Pendekatan ini telah membuktikan bahawa ia adalah putaran tidak varian dan juga dikenalpasti berupaya mengenalpasti dan memperoleh kewujudan imej sebagai sebahagian (subset) daripada satu imej.

CHAPTER 1 INTRODUCTION

Image retrieval attracts interest among researchers in the fields of image processing, multimedia, digital libraries, remote sensing, astronomy, database applications and others related area. An effective image retrieval system is able to operate on the collection of images to retrieve the relevant images based on the query image which conforms as closely as possible to human perception.

A major class of users' requests require retrieving images in the database that are spatially similar to the query image *(Manmatha and Ravela, 1997)*. Therefore, spatial information has been addressed as a good description and representation of an image, it able to capture meaningful contents of the image. Spatial information needs to be integrated with visual features to effectively retrieve similar images.

1.1 Background

Image retrieval has been an active research area since 1970s, with the thrust from 2 major research communities, database management and computer vision *(Yong et al, 1999)*. These two research communities study image retrieval from different perspectives, one being text-based and the other visual based *(Yong et al, 1999)*.

Text-based image retrieval techniques employ text to describe the content of the image which often causes ambiguity and inadequacy in performing an image database search and query processing. This problem is due to the difficulty to specify exact terms and phrases in describing the content of images. Since the textual annotations is based on language, variations in annotation will pose challenges to image retrieval.

To overcome this problem, content-based image retrieval (CBIR) has been used as an alternative to text based image retrieval. The general structure of content based image retrieval framework is shown in Figure 1.1. CBIR can be categorized based on the type of features used for retrieval which could be either low level or high level features. Low level features include colour, texture and shape. High level features are more abstract, whereby one or more low level features could be combined together to form a high level feature. Examples of high level features include concept, object and others. Users normally specify high level abstract features for images to be retrieved. These high level features need to be translated into visual features before retrieval can be performed. For example, a high level query can be on images that have red colour of a flower with a star shape and block texture.

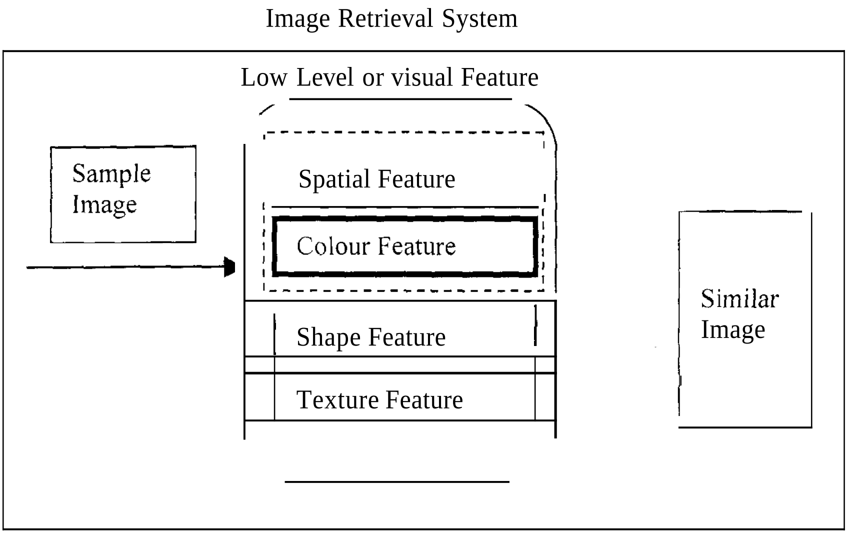

Figure 1.1 Content-Based Image Retrieval Framework

This research is based on low level features emphasizing and highlighting colour-spatial based features in image retrieval. By just using a single low visual features such as colour, texture or shape alone. it is not sufficient to produce satisfactory and accurate results *(Wang,* 2000). As spatial information is able to capture meaningful contents of images. it is generally accepted as an important feature to be integrated with visual features for image retrieval.

Examples of Colour-spatial approaches include *Chitkara et at, 2000; Chua et al, 1997; Cinque et al, 1999; Jo and Um, 2000, Kankanhalli et al, 1997; Wang,* 2000. An effective colour-spatial approach is needed to achieve better accuracy, which effectively conform to human perception.

The general structure of the image retrieval system is shown in Figure 1.2. It consists of two modules which are the query handler module and database builder module. The query handler provides a sample image and the system retrieves the most similar images based on a given query image. The Database builder module is used to store and manage the images. Feature extraction process enables of extracting the features for either the query handler module or the database builder module. Similarity measurement is used to calculate the distance for all the images in database relative to the given query image.

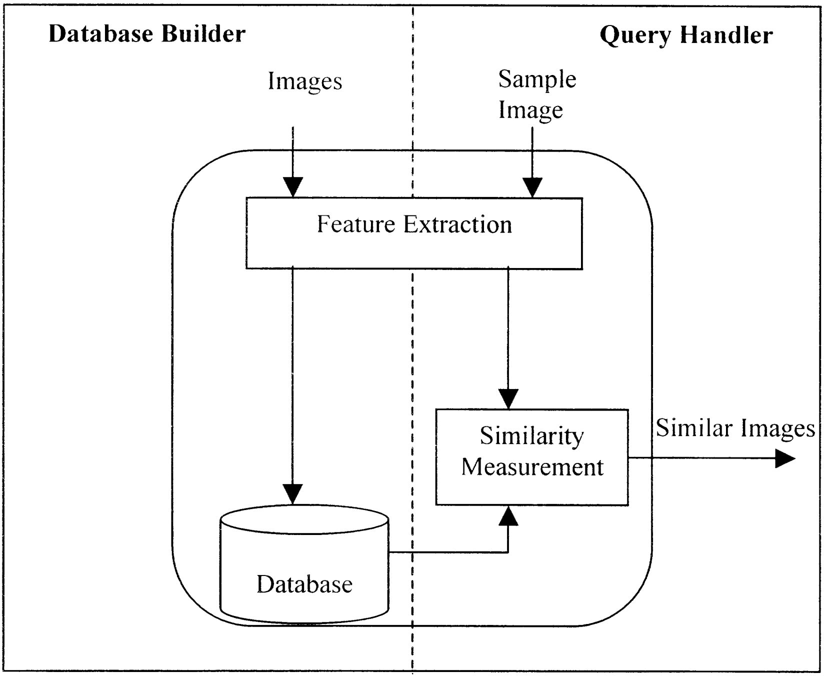

Figure 1.2 General Structure Of The Image Retrieval System

1.2 Problem Statement

Similarity based retrieval of images is an important task in many image database applications. By identifying and mapping an image based on colour, texture or shape alone is not sufficient to produce satisfactory and accurate results. Although the amount of each colour, shape and texture entry in an image may be similar to another image, but these images may appear to be visually and physically different. For example, consider two images, one of a red flower and the other a red t-shirt. Even though these two images are entirely different, both could be similar in terms of colour composition and could be retrieved as similar images if the image retrieval scheme is based only on colour. On the other hand, when comparing the colour clustering and positions of the clusters, these images can be very different.

According *Ooi et al. (1998)* problems *with* major image retrieval schemes are due to the lack of a good measurement of visual similarity, the importance accorded to use interaction and feedback, and the neglect of spatial information. Spatial feature thus can be used to capture meaningful contents of the image by identifying the layout attribute, arrangement and distribution of colour.

In order to meet the major class of users' requests, some researcher's have integrated colour and spatial features (*Chitkara et at, 2000; Chua et al, 1997; Cinque et al, 1999; Jo and Um, 2000; Kankanhalli et al, 1997; Wang, 2000).* Addressing transform invariance of images has been a problem in a number of approaches. Shifted or translation invariance has been addressed by all of the colour-spatial approaches mentioned. Significant problems include the failure to locate partial images and inability to detect rotations and feature scaling in image retrieval. The approaches (*Chitkara et at, 2000; Chua et al, 1997, Cinque* ***et*** *al, 1999; Jo and Um, 2000; Kankanhalli* ***et*** *al. 1997; Wang,* 2000) tend to be rotation invariant up to limited set of degrees of rotation and certain classes of images. However. all approaches mentioned did not address the ability to locate and retrieve partial images.

In this research, a system combining the spatial and colour features is proposed to solve the existing problems mentioned.

1.3 Objectives

The objectives of the research are as follows:

- To study current colour-spatial similarity based image retrieval systems.
- To investigate and address the problems of current techniques of colour-spatial similarity based images retrieval systems.
- To study and model spatial feature extraction and representation.
- To produce and design the system architecture by integrating spatial and colour features to handle existing problems and improve the accuracy of similarity measurement for image retrieval.
- To compare and contrast the performance of proposed approach.

1.4 Significance Of The Research

We evaluate feature and spatial queries to demonstrate the improved image query capabilities over existing content-based approaches. The outcome would explore the similarity measurement technique for the images by incorporating spatial feature in the similarity measurement process. With the incorporation of spatial and visual features paradigmwe aim to create a more effective, efficient and accurate system for retrieving colour images that conform to human perception.

1.5 Scope

This research focuses mainly on the study of suitability and usability of spatial features for image retrieval. The research explores on the integration of spatial feature and colour feature to demonstrate the improvement in terms of retrieval accuracy. The scope of this research will be focused on achieving rotation invariant and partial-images detection.

1.6 Research Methodology

The research is divided into three stages; the first stage is to survey and review related research in the area of colour and spatial similarity measurement. We investigate and address the problems of the current techniques of the spatial similarity based retrieval. This stage studies the requirements and specifications for the colour features and the spatial similarity measurement. The second stage concentrates on designing a solution for integrating spatial information with colour features in an image retrieval system. The final stage is to develop a prototype system according to the design proposed. Evaluations are conducted to test the efficiency and effectiveness of the proposed solutions.

The three stages mentioned can be summarized as below :

Literature Review

- Review on related work done
- Investigate and address the problems of the current techniques on the spatial based images retrieval.
- Do a comparison of the different methods used for spatial similarity measurement methods or functions.
- Identify the specifications and requirements of spatial information integration with colour features

Design The Solution

- Design and model the spatial information for image retrieval
- Propose a design for Integrating spatial and colour features for image similarity measurement

Development and Experiments

- Implement the proposed solution by developing a prototype system for image retrieval
- Evaluate the performance of the prototype and compare the performance
- Analyze and refine the proposed solution.

1.7 Thesis Organization

The structure of the thesis is organized in the following order :

Chapter 2 : The Literature Review chapter discusses the current existing techniques for colour-spatial based image retrieval by analyzing the image feature extraction techniques, image feature representation and also on similarity measurement. The comparison and discussion of each of this technique is discussed.

Chapter 3 : The System Design chapter describes the approach towards solving the problems highlighted earlier. The design of our proposed system including the system architecture, data flows diagram, feature extraction techniques, feature representation and the similarity measurement are covered in this chapter.

Chapter 4 : The Prototype System chapter presents our implementation tools, main algorithms employed, data structures, database structures as well as user interface for the develop prototype system.

Chapter 5 : Presents the experiments and the results of the developed system. An evaluation of the performance of the developed system is then discussed.

Chapter 6 : The conclusion chapter summarizes the research work and concludes the achievement of this research, as well as the difficulties and system limitations encountered. Future works are then presented.

CHAPTER 2 LITERATURE REVIEW

This chapter focuses on the study of various existing colour-spatial approaches. The various approaches reviewed, can be categorized into 3 categories: signature based approach, histogram approach and colour cluster approach. For each of the approaches, the features that are usually used, the representations of these features and similarity measurement are reviewed and covered.

The study on the strengths and the weaknesses of each approach is also discussed and summarized. The suitability and effectiveness of the existing colour-spatial is also highlighted.

2.1 Colour-Spatial Retrieval Technique

This section discusses our study of existing colour-spatial retrieval techniques.

2.1.1 Signature Based Approach

In this approach, bit streams are used to represent images. The existence of associated colours for each sub area in the image is then determined.

Chua et al. (1997) carried out their signature based approach by first constructing the colour histogram of the image and dominant colours are then chosen as colour representation of the image. The image is then partitioned into m x n cells of equal size to get the colour and spatial distribution. For every single colour, each cell is examined to determine the percentage of the total number of pixels in the cell is greater than a pre-defined threshold value. The cell is represented by a bit 1 if the cell satisfies the threshold value otherwise the cell is represented by a bit value 0.

Features for the each single colour, in bit stream for image i, F(i) is represented as

$$F(i) = \{b_0, b_1, b_2 \ldots\ldots b_{31}\}$$

Image Representation of the image

where b_0, b_1, $b_2 \ldots\ldots b_{31}$ represent the bit streams of the cell for the partitioned image in 32 bit signature.

For the similarity function, Q_I and D_I denote the signature of colour i for a query image Q and a database image D respectively. Let the representative colour sets of Q and D be C_Q and C_D respectively. Then the similarity measure, $S(Q,D,i)$, between Q and D for a colour $I \in C_Q$ can be determined as :

$$S(Q,D,i) = \sum_{j \in S_p} \frac{BitSet\ (Q_i \wedge D_j)}{BitSet\ (Q_i)} x SM(i,j)$$

Similarity Function with Perceptually Similar Colours

where *BitSet(BS)* denotes the number of bits in the bit stream *BS* that are set and $\wedge$ represents the bitwise logical-AND operation.

This computation also considers the contribution of the perceptually similar colour to that of colour i. S_p is the set of colours that is perceptually similar to colour i as derived from the colour similarity matrix *SM*. *SM(i,j)* denotes the *(i,j)* entry of matrix *SM*. The similarity measure between two image Q and D is the sum of the similarity for each colour in the representative set C_Q for image Q :

$$S(Q,D) = \sum_{\forall i \in C_Q} S(Q,D,i)$$

Similarity Function for Query Image Q and Database Image D

Usually all the dominant colours of the image are retrieved with the same weight. However, in certain applications, it is desirable to retrieve objects instead of the background. The weighted similarity function is applied and a higher weight is given to the particular object colour. The weighted similarity function is given as below,

$$S_{weight}(Q,D) = \sum_{i \in C_i} S(Q,D,i) + wt \sum_{i \in C_c} s(Q,D,i)$$

Weighted Similarity Function

where C_i and C_c are the set of background and object colours of Q respectively, and (wt > 1) is the weightage given to the object colours. A weight value greater than 1 can be assigned to the object colours to assign a higher weightage to images with similar object colour to that of the query image.

Jo and Um (2000) produce better results than *Chua et al's* approach by proposing a two axial signature-based method. Similar to *Chua et al's* approach, a histogram is constructed and the dominant colour is chosen to be the colour representative of the image. The dominant colour is represented by bit streams in discrete colour bins which acts as a filter for the image retrieval process.

The image is portioned into a grid of $n \times n$ subregions of equal sizes and is used to capture the colour-spatial information. Two forms of feature extraction are performed in this method, which are the Dominant Colour Composition (DCC) signature extraction and Dominant Colour Distribution (DCD) signature feature extraction. DCC signature is created with regionally representative colour set of an image. The DCC signature format is shown as below :

$$DCC = g,\ h_0,\ h_1 \ldots\ h_{16},\ h_{17}$$

DCC signature format

where g is the gray bin and the h_0, h_1 ... are for the hue bins. The size of the DCC signature is 19 bits.

The Dominant Colour Distribution (DCD) signature feature extraction is used to represent the spatial information of the dominant colour. The DCD signature consists of two sub signature namely DCD_h and DCD_v. Both DCD_h and DCD_v signatures are based on the horizontal and vertical axis of an image where dominant coloured regions are positioned. It can be expanded into a subregion with a colour C_k represented by bits along its horizontal and vertical axis respectively to get the zoned signature.

Bit streams are used to represent the DCD_h and DCD_v signatures. Let b_x and b_y be a bit position of bit stream such as $b_n, b_1, \ldots, b_{n-2}, b_{n-1}$ for horizontal and vertical axis respectively. b_n is set to bit value 1 if a dominant colour *exists* in that coordinate system in the horizontal or vertical axis otherwise the bit value is set to 0.

$$DCD_{Ck} = \{h_0, h_1 \ldots h_{n-1}, V_0, V_1 \ldots V_{n-1}\}$$

DCD Signature of the dominant colour C_k

DCD_{Ck} is the DCD signature of image i with the dominant colour k where $h_0, h_1 \ldots h_{n-1}$ are the bits represented in horizontal axis and $V_0, V_1 \ldots V_{n-1}$ are the bits represented in the vertical axis.

The distance function of the DCC signature that is used as a filter to prune away unrelated images works as below :

Let DCC_Q and DCC_I denotes the DCC signature for query image I and image Q in database. Distance measure for DCC is computed as below:

$$d_{DCC}(Q, I) = \frac{OnBit\ (DCC_Q \wedge DCC_I)}{OnBit\ (DCC_Q)}$$

Distance Measure for DCC signature

C_Q denotes the set of dominant colours of query image Q, and C_i denotes the set of dominant colours of a target image I. There exists $C_i \in C_Q$ and $C_j \in C_I$, where $0 \leq I, j \leq m-1$. The colour similarity is used to estimate the similarity of the colour between 2 images (query image and database image). The colour similarity function is measured as below:

$$d_C(Q,I) = \frac{1}{N} \sum_{C_i \in C_Q, C_j \in C_I} \max\ (d_{hsv}(C_i, C_j))$$

Colour Similarity Function

where N is the number of C_i for normalizing the distance value within [0,1]

The Spatial similarity or the distribution of the colour between 2 images is considered as similar if and only if there exists some bits overlapped between DCD_{Ci} and DCD_{Cj}. The spatial similarity function is given as below:

$$d_{DCD}(C_i, C_j) = \frac{OnBit\ (DCD_{Ci} \wedge DCD_{Cj})}{OnBit\ (DCD_{Ci})}$$

Spatial similarity function for each colour pair

The overall spatial similarity is evaluated as the function below,

$$d_{DCD}(Q,I)\ \frac{1}{N} \sum_{C_i \in C_Q, C_j \in C_I} d_{DCD}(C_i, C_j).$$

Overall spatial similarity function

where N is the number of C_i for normalizing the distance value within [0,1]

Finally, the image similarity is evaluated as below :

$$d(Q,I) = d_{DCD}(Q,I)\alpha + (Q,I)\beta + (Q,I)$$

where γ, α,β,γ are weight values within [0,1] and $\alpha+\beta+\gamma = 1$.

Chitkara et al.(2000) produce a signature based algorithms that is based on the compact representation of the information in a global colour histogram by decomposing the colour distributions of an image into bins that accommodate varying percentage compositions.

The image is then partitioned into fixed number of n representative colours $C=(c_1, c_2\ldots,c_n)$ Each colour c_j is then discretized into t binary bins $B^j = b_1^j, b_2^j \ldots\ldots b_t^j$ of equal or varying capacities of bin sizes. Equal capacity of bins is referred to as Constant Bin Allocation (CBA) and the varying capabilities of bins then is referred to as Variable Bin allocation (VBA). The percentages of each colour exists in the images is calculated and assigned to bins that accommodate varying percentage compositions. The bins are set accordingly as below :

$$b_i^j = \begin{cases} 1 \text{ if } I = [h_j \times t] \\ 0 \text{ otherwise} \end{cases}$$

where h_j represents the percentage pixel dominance of colour c_j and t is the total bins in image.

For an image Q with n colours and t bins, the bit signature of this image would then be represented by the following bit-string :

$$S_1 = B_I^1, B_I^2 \ldots. B_I^n$$
$$S = b_1^1 b_2^1 \ldots b_t^1 b_1^2 b_2^2 \ldots. b_t^2 \ldots. b_1^n b_2^n \ldots. b_t^n$$

Image representation of the image

The image is partitioned into cells to take the spatial information into consideration. Each cell of the image is then computed to get the bit signature and cell by cell comparison among two images computed using the similarity function to get the cumulative distance among such cells.

Similarity between the images can be computed as below:

$$d_0(Q,I) = \sum_{j=1}^{n} |(pos(B_Q^j) - pos(B_I^j)|^2$$

Image Similarity Function

where pos(B_R^K) gives the position of the set bits within B^k of image R. The bigger the d_0 value, the smaller the similarity value between images.

2.1.2 Spatial Chromatic Histogram Approach

Cinque et al. (1999)'s algorithms can extract and capture the information of the location of the pixels having the same colours and spatial arrangement of an image.

The histogram of the image is constructed in order to get the colour representation of the image. The image is then partitioned into n x m regions to represent the image using the coordinate system. $A_k^I := \{(x,y) \in I : I[x,y] = k\}$ as the set of pixels in the region of the image having the same colour k then is extracted. *H(k)* as the ratio of pixels having colour k in image I is obtained as below

$$H_I(k) = \frac{|A_k^I|}{n \bullet m}$$

Baricenter of the colour k in image, $b_I(k) := (\overline{x_k}, \overline{y_k})$ is obtained to give an idea of the position of the pixels having the same colour that can be computed as below

$$\overline{x_k} = \frac{1}{n} \frac{1}{|A_k|} \sum_{(x,y)\in A_k} x$$

$$\overline{y_k} = \frac{1}{m} \frac{1}{|A_k|} \sum_{(x,y)\in A_k} y$$

Baricenter of the pixels in A_k^I

Standard deviation of pixels in A_k^I, with relative coordinates (x_p, y_p) are then obtained to get the spatial distribution of the colours as below :

$$\sigma_I(k) = \sqrt{\frac{1}{|A_k|} \sum_{p\in A_k} d(p, b_I(k))^2}$$

Standard deviation of the pixels in A_k^I