ISSN: 2319-8753

International Journal of Innovative Research in Science, Engineering and Technology
Vol. 2, Issue 5, May 2013

IMAGE MATCHING ALGORITHM BASED ON HUMAN PERCEPTION

Mrigank Sharma [1], Amrita Priyam [2]

Final Semester Student, MCA, Department of Computer Science & Engineering, BIT, Ranchi, India[1],

Associate Professor, Department of Computer Science & Engineering, BIT, Ranchi, India[2]

Abstract: There has been in depth analysis and numerous researches on image matching and storage. Major software distributors providing RDBMS provide image storage based on textual description. Since Images can easily be searched based on textual information, it is not so efficient and hence brings need of such technique that allows search based on image feature and not based on textual description of image. In recent years dramatic changes have been seen in digital image libraries and other multimedia databases. In order to effectively and precisely retrieve the desired images from a large image database, the development of a content-based image matching system has become an important research issue. However, most of the proposed approaches emphasize on finding the best representation for different image features. Human perception of comparing image is something else. Visual perception is the ability to interpret the surrounding environment by processing information that is contained in visible light. The resulting perception is also known as eyesight, sight, or vision. Hence color attributes like the mean value, the standard deviation, and the image bitmap of a color image are used as the features for matching. In addition, the entropy based on the gray level co-occurrence matrix and the edge histogram of an image is also considered as the texture features.

Keywords: Image Processing, Human Perception, Pixel Generalisation

I. INTRODUCTION

Visual perception is the ability to interpret the surrounding environment by processing information that is contained in visible light. The resulting perception is also known as eyesight, sight, or vision. The various physiological components involved in vision are referred to collectively as the visual system, and are the focus of much research in psychology, cognitive science, neuroscience, and molecular biology.

There is some evidence (including disorders such as prosopagnosia) that face recognition is distinct from object recognition in terms of visual processing. For example, newborns show a preference for following moving faces within the first 30 minutes of life. However, some studies have shown that visual processing of complex non-face shapes happens in the same area of the brain as facial recognition. This implies it may be complexity, rather than the face per se, that influences visual processing in a distinct way.

Theories and observations of visual perception have been the main source of inspiration for computer vision (also called machine vision, or computational vision). Special hardware structures and software algorithms provide machines with the capability to interpret the images coming from a camera or a sensor. Artificial Visual Perception has long been used in the industry and is now entering the domains of automotive and robotics.

Rapid advances in science and technology have produced a large amount of image data in diverse areas, such as entertainment, art galleries, fashion design, education, medicine, industry, etc. We often need to efficiently store and retrieve image data to perform assigned tasks and to make a decision. Therefore, developing proper tools for the matching image from large image collections that matches image perception of human eye and the way human compare two images is challenging.

Two different types of approaches, i.e., text- and content-based, are usually adopted in image matching. In the text-based system, the images are manually annotated by text descriptors and then used by a database management system to perform image matching. However, there are two limitations of using keywords to achieve image matching: the vast amount of labour required in manual image annotation and the task of describing image content is highly subjective. That is, the perspective of textual descriptions given by an annotator could be different from the perspective of a user. In other words, there are inconsistencies between user textual queries and image annotations or descriptions. To alleviate the inconsistency problem, the image matching is carried out according to the image contents. Such strategy is

ISSN: 2319-8753

International Journal of Innovative Research in Science, Engineering and Technology
Vol. 2, Issue 5, May 2013

the so-called content-based image retrieval (CBIR). The primary goal of the CBIR system is to construct meaningful descriptions of physical attributes from images to facilitate efficient and effective matching *[1], [2]*.

CBIR has become an active and fast-advancing research area in image matching in the last decade. By and large, research activities in CBIR have progressed in four major directions: global image properties based, region level features based, relevance feedback, and semantic based. Initially, developed algorithms exploit the low-level features of the image such as color, texture, and shape of an object to help retrieve images. They are easy to implement and perform well for images that are either simple or contain few semantic contents. However, the semantics of an image are difficult to be revealed by the visual features, and these algorithms have many limitations when dealing with broad content image database. Therefore, in order to improve the matching accuracy of CBIR systems, region-based image matching methods via image segmentation were introduced. These methods attempt to overcome the drawbacks of global features by representing images at object level, which is intended to be close to the perception of human visual system. However, the performance of these methods mainly relies on the results of segmentation.

The difference between the user's information need and the image representation is called the semantic gap in CBIR systems. The limited matching accuracy of image centric matching systems is essentially due to the inherent semantic gap. In order to reduce the gap, the interactive relevance feedback is introduced into CBIR. The basic idea behind relevance feedback is to incorporate human perception subjectivity into the query process and provide users with the opportunity to evaluate the matching results. The similarity measures are automatically refined on the basis of these evaluations. However, although relevance feedback can significantly improve the matching performance, its applicability still suffers from a few drawbacks *[3]*. The semantic-based image matching methods try to discover the real semantic meaning of an image and use it to retrieve relevant images. However, understanding and discovering the semantics of a piece of information are high-level cognitive tasks and thus hard to automate.

A wide variety of CBIR algorithms has been proposed, but most of them focus on the similarity computation phase to efficiently find a specific image or a group of images that are similar to the given query. In order to achieve a better approximation of the user's information need for the following search in the image database, involving user's interaction is necessary for a CBIR system. In this paper, we propose a user-oriented CBIR system that uses the interactive algorithm *[6]* to infer which images in the data-bases would be of most interest to the user. Three visual features, color, texture, and edge, of an image are utilized in our approach. Algorithm provides an interactive mechanism to better capture user's intention. There are very few CBIR systems considering human's knowledge, but is the representative one. They considered the red, green, and blue (RGB) color model and wavelet coefficients to extract image features. In their system, the query procedure is based on association (e.g., the user browses an image collection to choose the most suitable ones). The main properties of this paper that are different from it can be identified as follows:

1) Image features— color features from the hue, saturation, value (HSV) color space, as well as texture and edge descriptors, are adopted in our approach and,
2) Search technique —our system adopts the query-by-example strategy (i.e., the user provides an image query). In addition, we hybrid the user's subjective evaluation and intrinsic characteristics of the images in the image matching against only considering human judgment.

II. METHODOLOGY

One of the key issues in querying image databases by similarity is the choice of appropriate image descriptors and corresponding similarity measures. In this section, we first present a brief review of considered low-level visual features in our approach and then review the basic concept of the algorithm.

A color image can be represented using three primaries of a color space. Since the RGB space does not correspond to the human way of perceiving the colors and does not separate the luminance component from the chrominance ones, we used the HSV color space in our approach. HSV is an intuitive color space in the sense that each component contributes directly to visual perception, and it is common for image matching systems *[9], [17]*. Hue is used to distinguish colors, whereas saturation gives a measure of the percentage of white light added to a pure color. Value refers to the perceived light intensity. The important advantages of HSV color space are as follows: good compatibility with human intuition, separability of chromatic and achromatic components, and possibility of preferring one component to other *[18]*.
The color distribution of pixels in an image contains sufficient information. The mean of pixel colors states the principal color of the image, and the standard deviation of pixel colors can depict the variation of pixel colors. The variation degree of pixel colors in an image is called the color complexity of the image. We can use these two features

ISSN: 2319-8753

International Journal of Innovative Research in Science, Engineering and Technology
Vol. 2, Issue 5, May 2013

to represent the global properties of an image. The mean(μ)and the standard deviation(σ)of a color image are defined as follows:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} P_i \qquad (1)$$

$$\sigma = \left[\frac{1}{N-1}\sum_{i=1}^{N}(P_i - \mu)^2\right]^{1/2} \qquad (2)$$

Where μ = [μH, μS, μV] and σ = [σH, σS, σV] , each component of μ and σ indicates the HSV information, respectively, and Pi indicates the i pixel of an image. In addition to the global property of an image, the local color properties in an image play also an important role to improve the matching performance. Hence, a feature called binary bitmap can be used to capture the local color information of an image. The basic concept of binary bitmap comes from the block truncation coding *[19]*, which is a relatively simple image coding technique and has been successfully employed in many image processing applications. There are three steps to generate the image binary bitmap. This method first divides an image into several non overlapping blocks. Let Bj = {b1,b2,...,bk} be the jth block of the image, where 1≤j≤m; k represents the total number of pixels in the block, and m is the total number of blocks in the image. The second step is to compute the mean value for each block. Let μBj be the mean value of the block Bj, which is defined as follows:

$$\mu_{B_j} = \frac{1}{k}\sum_{i=1}^{\mathrm{k}} \mathrm{B}_i \qquad (3)$$

Where μBj = [μHBj,μSBj,μVBj]T. In the final step, comparing μBj with the image mean value(μ) is performed to determine the characteristic of the block Bj and to generate the image binary bitmap. Hence, suppose that I=[IH, IS, IV] is the binary bitmap of the given image. Each component in I is expressed as IH=[IH1,IH2,...,IHm], IS=[IS1,IS2,...,ISm], and IV=[IV1,IV2,...,IVm], respectively. The entries are represented by

$$IH_j = \begin{cases} 1, & if\ \mu H_{B_j} \geq \mu H \\ 0, & otherwise \end{cases} \qquad (4)$$

$$IS_j = \begin{cases} 1, & if\ \mu S_{B_j} \geq \mu S \\ 0, & otherwise \end{cases} \qquad (5)$$

$$IV_j = \begin{cases} 1, & if\ \mu V_{B_j} \geq \mu V \\ 0, & otherwise \end{cases} \qquad (6)$$

Texture is an important attribute that refers to innate surface properties of an object and their relationship to the surrounding environment. If we could choose appropriate texture descriptors, the performance of the CBIR should be improved. We use a gray level co-occurrence matrix (GLCM), which is a simple and effective method for representing texture *[20]*. The GLCM represents the probability p(i, j;d, θ) that two pixels in an image, which are located with distance d and angle θ, have gray levels i and j. The GLCM is mathematically defined as follows:

$$p(i,j;d,\theta) = \#\left\{(x_1,y_1)(x_2,y_2) | g(x_1,y_1) = i, g(x_2,y_2) = j,\ |(x_1,y_1) - (x_2,y_2)| = d, \angle\left((x_1,y_1),(x_2,y_2)\right) = \theta\ \right\} \qquad (7)$$

Where # denotes the number of occurrences inside the window, with I and j being the intensity levels of the first pixel and the second pixel at positions(x1,y1)and(x2,y2), respectively.
In order to simplify and reduce the computation effort, we computed the GLCM according to one direction (i.e., θ=0◦) with a given distance d (= 1) and calculated the entropy, which is used most frequently in the literature. The entropy (E) is used to capture the textural information in an image and is defined as follows:

$$E = -\sum_{i,j} C_{i,j} \log C_{i,j} \qquad (8)$$

Where C i,j is the GLCM. Entropy gives a measure of complexity of the image. Complex textures tend to have higher entropy.

Edges in images constitute an important feature to represent their content. Human eyes are sensitive to edge features for image perception. One way of representing such an important edge feature is to use a histogram. An edge histogram in the image space represents the frequency and the directionality of the brightness changes in the image. We adopt the edge histogram descriptor (EHD) *[21]* to describe edge distribution with a histogram based on local edge distribution in an image. The extraction process of EHD consists of the following stages:

1) An image is divided into 4×4 sub images.
2) Each sub image is further partitioned into non overlapping image blocks with a small size.
3) The edges in each image block are categorized into five types: vertical, horizontal, 45º diagonal, 135º diagonal, and non directional edges.
4) Thus, the histogram for each sub image represents the relative frequency of occurrence of the five types of edges in the corresponding sub image.
5) After examining all image blocks in the sub image, the five-bin values are normalized by the total number of blocks in the sub image. Finally, the normalized bin values are quantized for the binary representation. These normalized and quantized bins constitute the EHD.

III. PROPOSED SYSTEM

In general, an image matching system usually provides a user interface for communicating with the user. It collects the required information from the user and displays the matching results to him. However, as the images are matched based on low-level description features, the target or the similar images may be far away from the query in the feature space, and they are not returned in the limited number of retrieved images of the first display. Therefore, in some matching systems, there is a relevance input from the user, where human and computer can interact to increase matching performance.

According to the aforementioned concept, we design a user-oriented image matching system. Our system operates in three phases.

1) Querying: The user provides a sample image as the query for the system.

2) Similarity computation: The system computes the similarity between the query image and the database images according to the aforementioned visual features and graphically represents the computed mean and standard deviation which represents principal color and the variation of pixel colors respectively.

3) Matching: The system retrieves and presents a sequence of images ranked in decreasing order of similarity. As a result, the user is able to find relevant images by getting the top-ranked images first.

IV. SYSTEM DEMONSTRATION

At first, we give an example to illustrate the practicability of our proposed system. A user submits an image containing a car as the query image into the system, and then, the similarity measurement module of the system compares the query features with those images in the database and finds the most similar images to the query image. These images are ranked based on the similarity. Under each image, a ranking percentage value is attached so that the user can see which images are relevant or irrelevant. The amount of percentage represents the degree of relevance. Higher percentage values higher the relevance.

ISSN: 2319-8753
International Journal of Innovative Research in Science, Engineering and Technology
Vol. 2, Issue 5, May 2013

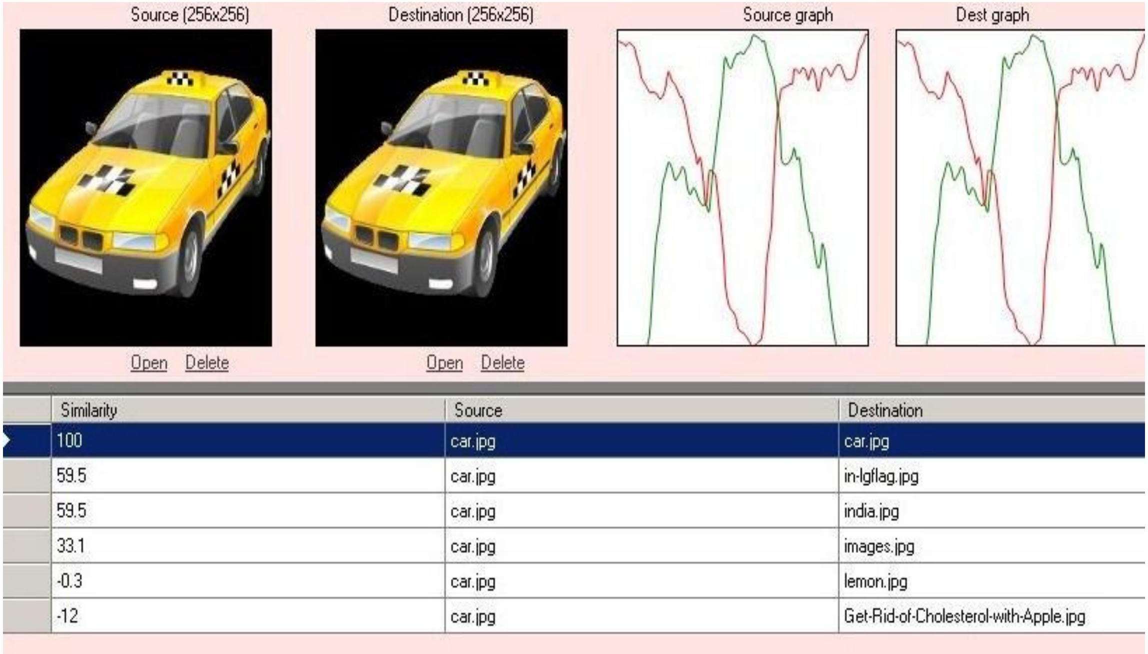

Fig. 1 A sample output Screen with two images matching with graph

The user can check all matches until he/she is satisfied with the matching results. Above figure shows the first display of returned images and the ranking list. The results are ranked in ascending order of similarity to the query image from top to bottom. Graphical representation shows mean of pixel colors and standard deviation of pixel colors which respectively states the principal color of the image, and the variation of pixel colors. Since we generalise full image through mean and standard deviation, graph helps to show the difference and due to which ranking of image relevance is pretty easier and visually shows the difference.

V. CONCLUSION

This paper has presented a user-oriented framework in inter-active CBIR system. In contrast to conventional approaches that are based on visual features, our method provides an interactive mechanism to bridge the gap between the visual features and the human perception. The color distributions, the mean value, the standard deviation, and image bitmap are used as color information of an image. In addition, the entropy based on the GLCM and edge histogram are considered as texture descriptors to help characterize the images. In particular, the algorithm can be considered and used as a semi automated exploration tool with the help of a user that can navigate a complex universe of images. Experimental results of the proposed approach have shown the significant improvement in matching performance. Further work considering more low-level image descriptors or high-level semantics in the proposed approach is in progress.

REFERENCES

[1] M. Antonelli, S. G. Dellepiane, and M. Goccia, “Design and implemen-tation of Web-based systems for image segmentation and CBIR,” IEEE Trans. Instrum. Meas., vol. 55, no. 6, pp. 1869–1877, Dec. 2006.
[2] N. Jhanwar, S. Chaudhuri, G. Seetharaman, and B. Zavidovique, “Con-tent based image retrieval using motif cooccurrence matrix,” Image Vis. Comput., vol. 22, no. 14, pp. 1211–1220, Dec. 2004.
[3] J. Han, K. N. Ngan, M. Li, and H.-J. Zhang, “A memory learning frame-work for effective image retrieval,” IEEE Trans. Image Process., vol. 14, no. 4, pp. 511–524, Apr. 2005.
[4] H. Takagi, “Interactive evolutionary computation: Fusion of the capacities of EC optimization and human evaluation,”Proc. IEEE, vol. 89, no. 9, pp. 1275–1296, Sep. 2001.
[5] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, “A survey of content-based image retrieval with high-level semantics,” Pattern Recognit., vol. 40, no. 1, pp. 262–282, Jan. 2007.
[6] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, “Content-based image retrieval at the end of the early years,”IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 12, pp. 1349–1380, Dec. 2000.
[7] S. Antani, R. Kasturi, and R. Jain, “A survey of the use of pattern recogni-tion methods for abstraction, indexing and retrieval of images and video,” Pattern Recognit., vol. 35, no. 4, pp. 945–965, Apr. 2002.
[8] X. S. Zhou and T. S. Huang, “Relevance feedback in content-based image retrieval: Some recent advances,”Inf. Sci., vol. 148, no. 1–4, pp. 129–137, Dec. 2002.
[9] H.-W. Yoo, H.-S. Park, and D.-S. Jang, “Expert system for color image retrieval,” Expert Syst. Appl., vol. 28, no. 2, pp. 347–357, Feb. 2005.
[10] T.-C. Lu and C.-C. Chang, “Color image retrieval technique based on color features and image bitmap,” Inf. Process. Manage., vol. 43, no. 2, pp. 461–472, Mar. 2007.
[11] A. Vadivel, S. Sural, and A. K. Majumdar, “An integrated color and intensity co-occurrence matrix,” Pattern Recognit. Lett., vol. 28, no. 8, pp. 974–983, Jun. 2007.

[12] M. H. Pi, C. S. Tong, S. K. Choy, and H. Zhang, “A fast and effective model for wavelet sub band histograms and its application in texture image retrieval,” IEEE Trans. Image Process., vol. 15, no. 10, pp. 3078–3088, Oct. 2006.
[13] M. Kokare, P. K. Biswas, and B. N. Chatterji, “Texture image retrieval us-ing new rotated complex wavelet filters,”IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 35, no. 6, pp. 1168–1178, Dec. 2005.
[14] M. Pi and H. Li, “Fractal indexing with the joint statistical properties and its application in texture image retrieval,”IET Image Process.,vol.2, no. 4, pp. 218–230, Aug. 2008.
[15] S. Liapis and G. Tziritas, “Color and texture image retrieval using chromaticity histograms and wavelet frames,”IEEE Trans. Multimedia,vol.6, no. 5, pp. 676–686, Oct. 2004.
[16] Y. D. Chun, N. C. Kim, and I. H. Jang, “Content-based image retrieval using multi resolution color and texture features,” IEEE Trans. Multimedia, vol. 10, no. 6, pp. 1073–1084, Oct. 2008.
[17] H. Nezamabadi-pour and E. Kabir, “Image retrieval using histograms of unicolor and bicolor and directional changes in intensity gradient,” Pattern Recognit. Lett., vol. 25, no. 14, pp. 1547–1557, Oct. 2004.
[18] K. N. Plantniotis and A. N. Venetsanopoulos,Color Image Processing and Applications. Heidelberg, Germany: Springer-Verlag, 2000.
[19] E. J. Delp and O. R. Mitchell, “Image coding using block truncation coding,” IEEE Trans. Commun., vol. COM-27, no. 9, pp. 1335–1342, Sep. 1979.
[20] R. M. Haralick and L. G. Shapiro, Computer and Robot Vision: Volume I. Reading, MA: Addison-Wesley, 1992.
[21] T. Sikora, “The MPEG-7 visual standard for content description—An overview,”IEEE Trans. Circuits Syst. Video Technol., vol. 11, no. 6, pp. 696–702, Jun. 2001.