論 文
7-5-10

# 칼라 분포정보를 이용한 성능적 이미지 검색 평가
# (Evaluation of the Use of Color Distributionfor Image Search in Various Setup)

이용환(Yong-Hwan Lee)[1] 안효창(Hyochang Ahn)[1]
이상범(Sang-Burm Rhee)[2] 박진양(Jin-Yang Park)[3]

## 요 약

최근 대용량의 디지털 이미지가 제작되면서 멀티미디어 관련 기술에서 이미지 검색이 많은 관심의 대상이 되고 있다. 본 논문에서는 이미지 검색(Image Search)을 위한 가장 기본적인 요소인 이미지 색상에 칼라 분포 정보를 이용하고 다양한 요소에 따라 가중치를 부여한 칼라 기반의 검색 기술자(Descriptor)를 제안하였고 시뮬레이션을 통하여 제안 기술자의 성능을 평가하였다. 칼라 히스토그램을 통한 이미지 검색 기술자를 설계하는데 있어 칼라모델은 HSV를, 웨이블릿 변환 필터는 Daubechies 9/7을, 웨이블릿 분해 레벨은 2레벨을 적용하였을 때 가장 좋은 검색 효율성을 보였다. 또한 유사도 검색은 히스토그램 이차행렬(Quadratic Matrix)을 적용하여 보다 나은 성능을 얻었으나 유사도 검색 계산 시간에서 절대 차이값의 합(L1 Norm)을 사용하는 경우에 비해 20배 이상의 처리 시간이 소요되었다.

## ABSTRACT

Image Search is one of the most exciting and fast growing research areas in the filed of multimedia technology. This paper conducts an empirical evaluation of color descriptor that uses the information of color distribution in color images, which is the most basic element for image search. With the experimental results, we observe that in the top 10% of precision, HSV, Daubechies 9/7 and 2 level decomposition have little better than others. Also histogram quadratic metrics outperform the Minkowski form distance metrics in similarity measurements, but spend more than 20 in computational times.

논문접수 : 2006. 9. 1.
심사완료 : 2006. 10. 1.

1)정회원 : 단국대학교 전자컴퓨터공학과 박사과정
2)정회원 : 단국대학교 전자컴퓨터공학과 교수
3)정회원 : 인하공업전문대학 컴퓨터정보과 교수

# 1. Introduction

Large collections of digital images are being created in many areas of commerce, hospital, academia and web site. Many of these collections are the product of digitizing existing collections of analog photographs, drawings and prints. These images are converted to digital from and stored in image database for later use. Or sometimes these are the generation from the digital device such as digital camera. This leads to issues such as reducing storage space and querying the database, that is, getting the information we need from the database as accurately and quickly as possible [1,2]. The conventional and currently used method of image retrieval is searching for a keyword that would match the descriptive word assigned to the image by a human categorizer [3]. Recently research on CBIR (Content-based Image Retrieval) has received much attention, and is one of the most exciting and fastest growing research areas in the field of multimedia technology [1,4].CBIR, which is based on automatically extracted primitive visual features such as color, texture, shape, and even the spatial relationship among objects, has been employed since the 1990's [4].

This study compares empirically the retrieval performance of the weighted color distribution descriptor in a challenging problem of content-based retrieval of semantic image categories. In this paper, an empirical estimation of search scheme is performed with various parameters such as color space, wavelet filters, wavelet decomposition levels and distance metrics.

The remainder of this paper is organized as follow. In the next section, general image search algorithm, which is discussed by JPSearch Ad-hoc group, is presented. We describe the color descriptor based on color distribution in section 3. In section 4, we then present some experimental results of our image search scheme, and the last section draws conclusions from these experiments.

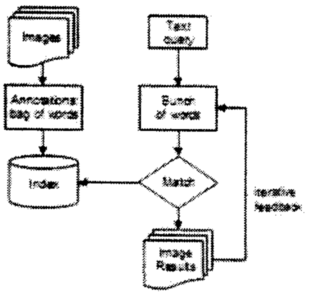
# 2. Related Literature on JPSearch

A lot of general-purpose image retrieval engines have been developed. We can not survey all related works in the allocated space. Thus, we try to emphasize come of the works that are most related to the standardization, especially mentioned by JPSearch AHG. This involved two steps [6]:

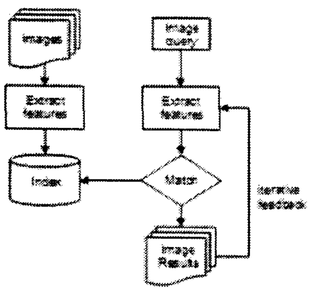
Step1. Extracting image features to a distinguishable extent;

Step2. Matching these features to yield a result that is visually similar;

Figure 1 shows the traditional approaches to digital image searching [2]. Image search system, shown in Figure 1(a), requires each image to be associated with one or more keywords entered by human. The second common form of image search system, Figure 1(b), uses an image as a query and attempts to retrieve other images which are similar. This is accepted state of the art in content-based image retrieval system from JPEG standardization JPSearch [2].




(a)




(b)

Figure 1. NaïveSystem View of Image Search and CBIR (a) Image Search and Retrieval using Annotation (e.g., keywords) (b) Image Search and Retrieval using an image as query (e.g., CBIR)

## 3. Weighted Color Distribution Extraction

Color is one of the most widely and extensively used visual features in the image retrieval [1,6]. This feature is relatively robust to background complication and independent of image size, resolution and orientation. Our image search scheme is depicted in Figure 2, which is used the scalable color histogram with weight for each channels.

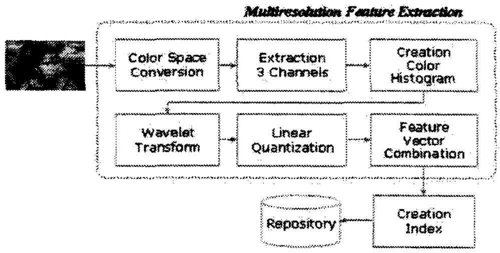



Figure 2. Diagram of Our Image Search Scheme

### 3.1 Weighted Color Histogram Descriptor

Descriptors for the color feature are mostly statistics of color distribution, average color and color moments [5]. Color histogram is one of the most frequently used color descriptor that characterizes the color distribution in an image [7]. WCHD (Weighted Color Histogram Descriptor) is compound descriptor consisting of color space, color histogram, quantization, wavelet transform and weighted combination of feature vectors. The step to extract the visual feature is the following.

---

Step1. Given the query image, convert the color space from original RGB to the desired color model;
Step2. Quantize the given image for processing efficiency;
Step3. Separate three color channels and compute the color histogram from the input image;
Step4. Perform the wavelet transform;
Step5. Linear quantization is performed to save the computation time;
Step6. Generate weighted combined feature vectors;
Step7. Create index and add it to image database;

---

The first step is color space conversion. We test three color models: HSV, CIELab, YCbCr to compare which is more suitable.

RGB is simple and easy to understand, but this is not uniform. Human's perception is more sensitive to green than red and blue, therefore the three color channels cannot be divided into the same scale [8]. So RGB is exception in this case. YCbCr is used in JPEG2000 standard compression, HSV is natural and intuitive to human for developing image processing algorithm based on color, and CIELab is the most uniform color space available [9,10]. The fourth step is wavelet transform. We estimate four wavelet filters, which are Haar, Db4, Daubechies 9/7 and 5/3. Haar is simple and efficient to computation time, and Daubechies 9/7 is used JPEG2000 standard [7,11]. Also we evaluate five level decompositions in wavelet transform. In sixth step, weight can be specified for each color channel and we give the default weight with 1, and then test the various factors.

3.2 Similarity Calculation

There are two functions for measuring similarity, one is Minkowski-form distance metrics and another is Quadratic-form distance metrics [12]. Minkowski-form distance, which compares only the same bins between color histograms, is as follows.

$$D(I_q, I_t) = \sum_{i=1}^{b} | H_q(i) - H_t(i) |^r \qquad (1)$$

where Iq and It are query image and target image, b is dimension of feature vector, Hq and Ht are the color histogram of query image and target image, respectively. When r=1, the distance metrics become L1, and when r=2, the distance metrics becomes the Euclidean distance. We use L1 norm distance between two color feature vectors in this paper.

Quadratic-form distance metrics, which compares not only the same bins but multiple bins between two histograms, is as follows [13].

$$D(I_q , I_t) = (H_q - H_t)^t \cdot A \cdot (H_q - H_t) \qquad (2)$$

where A=[aq,t] is a N*N matrix, N is the number of bins in the color histograms, and aq,t is calculated by Equation (3)

$$a_{q,t} = 1 - ( \sqrt{M_1^2 + M_2^2 + M_3^2} / \sqrt{5} ) \qquad (3)$$

where M1, M2 and M3 is below.

$M_1 = (H_q - H_t)^2$

$M_2 = (H_q(i) \cdot \sin(H_q(i)) - (H_t(j) \cdot \sin(H_t(j)))^2$

$M_3 = (H_q(i) \cdot \cos(H_q(i)) - (H_t(j) \cdot \cos(H_t(j)))^2$

## 4. Experimental Results

We implemented all approaches using Matlab 7.0.1. All the response times of our experiments were measured using a PC running Windows 2003 Server OS with a 1.8GHz Pentium IV CPU and 512Mbytes main memory.

4.1 Test Image Dataset

We have built on the image datasets obtained from producing some images ourselves and downloading on the website [14]. In this test image database, there are 1,000 different sizes (each of 10KB~700KB) and different resolutions (each of 640×480, 768×480 and 384×256) formatted JPEG images. The test images were manually partitioned into semantic categories by a human observer. A semantic category corresponds to a set of images, which he perceived to convey similar semantic meaning, not necessarily identical color content or spatial structure.

4.2 Retrieval Effectiveness

The most common evaluation measures used in IR (Information Retrieval) are recall and precision, usually presented as a Precision and Recall curve [15]. Precision is the probability of retrieved that relevant to query, and recall is the probability of relevant that are retrieved. Let A be the set of relevant images, B the set of retrieved images. Then recall and precision are defined as the following conditional probabilities [11,12].

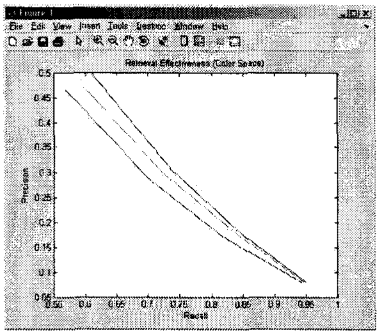
$$recall = P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{a}{a+c} \quad (4)$$

$$precision = P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{a}{a+b} \quad (5)$$

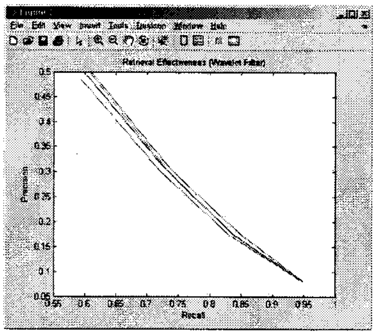
where a is the number of retrieved relevant images, b is the number of retrieved irrelevant images and c is the number of un-retrieved relevant images.
A semantic category is only used in calculating the recall and precision for the effectiveness. It is known whether any two images in the same category are similar to query results.

4.3 Results

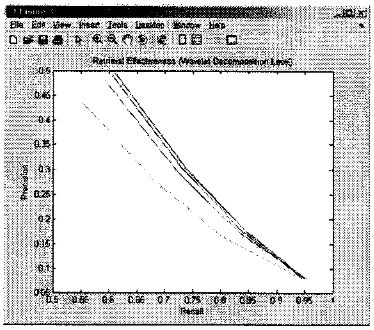
Each image of the twenty-eight semantic categories served as the query image in turn, i.e. 1,000 queries were performed in total. This way we obtain more reliable estimates on retrieval performance, in comparison to experiments where the results are based on a small number of queries, for example MPEG-7 Ground Truth Sets.
Figure 3shows retrieved results of ranking through 5%, 10%, 20% and 50% of whole image database with PR curve. Test image collection has 1,000 images, so then the numbers of results with ranking are 50, 100, 200 and 500, respectively. We observe that in the top 5% of recall, which is relevant in practical evaluation, HSV color space shows the best performance for retrieval under same other conditions (e.g., Daubechies 9/7 wavelet filter, 2 level decomposition and L1 as similarity measure function with same weight as each channels), shown in Figure 3(a). And the performance of wavelet filter ranks in the following order: Daubechies 9/7, 5/3, Db4 and Haar, but Haar is little better than others on computing time, shown in Figure 3(b). In the case of wavelet decomposition level, 2 level is best and others rank in the following order: 3,1,4,5, shown in Figure 3(c).




(a)



(b)



(c)

Figure 3. Retrieval Effective with PR Curve: (a) Comparison to color space on same other parameters (such as wavelet transform filter and similarity function) : red line is YCbCr, green is HSV and blue is RGB (b) Comparison to wavelet filter on same other parameters (such as color space and similarity function) : red is Dauchies9/7, green is 5/3 and blue is haar. (c) Comparison to wavelet decomposition level : red is 1, green is 2, blue is 3, magenta is 4 and cyan is 5 level decomposition.

Feature vector dimension and processing time are shown in Table 1 and the numeric results on the average time column are collected by an average of three times for each query image. Considering the difficulty of the problem, our image search scheme achieves a decent 35% average precision in retrieving the top 5% of the images from the semantic category of the query image.

Table 1. Image Search Results

| | | Average Processing Time [sec/image] | Retrieval Effectiveness with Top 5% |
|---|---|---|---|
| Extract Feature Vectors | | 0.24 | |
| Similarity Matching | L1 Norm | 0.45 | Recall 68% Precision 33% |
| | Quadratic Metrics | 25.9 | Recall 70% Precision 35% |

## 5. Conclusion and Future Works

Color is a feature of the great majority of content-based image retrieval systems. However, the robustness, effectiveness and efficiency of its use in image indexing are still open issues [1]. In this paper, our study compared the performance of weighted color distribution descriptor. The experimental results were carried out that the HSV, Daubechies 9/7, 2-level decomposition have little better than others at the color space, wavelet filter and wavelet decomposition level, respectively. And L1 metric is enough to use as a similarity measurement, compared with Quadratic distance metric, for computational reasons.

The experiments confirm that the image spatial organization of color is additional information to retrieve the similar. For this reason, some of spatial information is more necessary to decide whether color image is similar or not. This spatial information can be color layout descriptor, used in MPEG-7 Visual [7,11], or color auto-correlogram, which express how the spatial correlation of pairs of color changes with distance [16].

## References

[1] Raimondo Schettini, Gianluigi Ciocca, Silvia Zuffi, "A Survey of Methods for Colour Image Indexing and Retrieval in Image Databases", Color Imaging Science: Exploiting Digital Media, (R.Luo, L.MacDonald eds.), Wiley, 2001.

[2] Mun-Kew Leong, Wo Chang, ISO/IEC JTC1/SC29/ WG1N3684, "Framework and System Components", July, 2005.

[3] John R. Smith, Shih-Fu Chang, "Tools and Techniques for Color Image Retrieval", IS&T/SPIE, Symposium on Electronic Imaging Science and Technology - Storage and Retrieval for Image and Video Database IV, vol.2670, Feb., 1996.

[4] Arnold W.M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, Ramesh Jain, "Content-based Image Retrieval at the End of the Early Years", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.22, no.12, Dec., 2000.

[5] Zijun Yang, C. C. Jay Kuo, "Survey on Image Content Analysis, Indexing, and Retrieval Techniques and Status Report of MPEG-7", Tamkang Journal of Science and Engineering, vol.2, no.3, pp.101~118, 1999.

[6] Thomas Seidl, Hans-Peter Kriegel, "Efficient Use-Adaptable Similarity Search in Large Multimedia Database", Proc. Of the 23rd International Conference on Very Large Database VLDB, Athens, Greece, Aug., 1997.

[7] B. S. Manjunath, Jens-Rainer Ohm, Vinod V. Vasudevan, Akio Yamara, "Color and Texture Descriptor", IEEE Transactions on Circuits and Systems for Video Technology, vol.11, no.6, June, 2001.

[8] Linh Viet Tran, "Efficient Image Retrieval with Statistical Color Descriptors", Ph.D. Thesis, Dept. of Science and Technology, Kinkoping University, Sweden, 2003.

[9] Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins, "Digital Image Processing using MATLAB", Prentice Hall, 2004.

[10] Martin Boliek, Charilaois Christopoulos, Eric Majani: ISO/IEC JTC1/SC29/WG11/N1646R, "JPEG2000 Part I Final Committee Draft Version 1.0", April, 2000.

[11] Jose M. Martinez, ISO/IEC JTC1/SC29/WG11/N6828, "MPEG-7 Overview", Oct., 2004.

[12] Virrorio Castelli, Lawrence D. Bergman, "Image Databases: Search and Retrieval of Digital Imagery", Wiley Inter-Science, 2002.

[13] James Hafner, Harpreet S. Sawhney, Will Equits, Myron Flickner, Wayne Niblack, "Efficient Color Histogram Indexing for Quadratic form Distance Functions", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.17, no.7, July, 1995.

[14] Website: http://wang1.ist.psu.edu, http://muvis.cs.tut.fi, http://www.cs.washington.edu, ftp://ftp.sunet.se

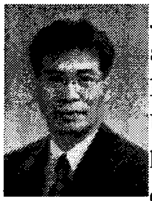
[15] Dionysius P. Huijsmans, Nicu Sebe, "How to Complete Performance Graphs in Content-based Image Retrieval: Add Generality and Normalize Scope", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.27, no.2, pp.245~251, Feb., 2005.

[16] Huang J, Kumar SR, Mitra M, Zhu HJ, "Image Indexing using Color Correlograms", Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.762~768, 1997.

Yong-Hwan Lee received the BS and MS degrees in Computer Science from Dankook Univ. in 1993 and 1995, respectively. During 1995~2003, he stayed in Korea Information System Inc. and eKalos Inc. to develop the ERP, EP, ITA solutions. And now he is undertaking a doctorate course as a member of the multimedia application lab at Dankook Univ. His research interests include Image and Video Processing, Image Search (JPSearch), Face Recognition, Standardization for Image and Video (JPEG and MPEG).

Hyochang Ahn received the BS and MS in Computer Science from Sangji Univ. in 2003 and Dankook Univ. in 2006, respectively. Now he is undertaking a doctorate course as a member of the multimedia application lab at Dankook Univ. His research interests include Image and Video Processing, Fingerprint Recognition, Face Recognition, and ARM Processor.

Sang-Burm Rhee received a Ph.D. degree in Electronics Engineering from Yonsei Univ. in 1986. He has been a professor at Dankook Univ. since 1979. His research interests are in the area of Microprocessors, SoC (System-On-Chip), Pattern Recognition, Multimedia Processing, Video/Audio Watermarking. They include topics such as Object-oriented Methods for Audio/Video Watermarking, Pattern Recognition and HDL for SoC.



Jin-Yang Park received BS, MS and Ph.D. degrees in Electronics Engineering from Dankook Univ. in 1982,1984 and 1997. He has been a professor at department of computer science Inha technical college since 1988. His research interests are in the area of Microprocessors, SoC (System-On-Chip), Pattern Recognition and algorithm for robot driving S/W.