
시각 특징과 퍼지 적분을 이용한 내용기반 영상 검색

Content-Based Image Retrieval Using Visual Features and Fuzzy Integral

송영준, 김동우, 김 남, 김미혜
충북대학교 전기전자컴퓨터공학부

Young-Jun Song(songyjorg@dreamwiz.com), Dong-Woo Kim(dubssi@paran.com),
Nam Kim(namkim@cbnu.ac.kr), Mi-Hye Kim(mhkim@cbnu.ac.kr)

요약

본 논문은 공간주파수 특징들과 다중 해상도 특징들을 가진 웨이블릿 영역에서 추출된 각 대역의 시각 특징 추출과 이들의 퍼지 적분 조합에 대하여 제안하였다. 칼라 양자화 이후에 똑같은 칼라의 빈도를 취함으로써 기존의 칼라 히스토그램 인터섹션 방법의 단점인 양자화 에러를 줄일 수 있게 칼라 특징을 표현한다. 또한 유사도는 서로 독립적인 특성을 갖는 호모그램, 칼라, 에너지 특징을 퍼지 측도와 퍼지 적분을 사용하여 조합한다. 1,000개의 칼라 영상에 대하여 실험을 하였고, 제안된 방법이 기존 방법들보다 객관적이고 주관적인 성능에서 우수함을 보였다.

■ 중심어 : | 내용기반검색 | 퍼지적분 | 특징 추출 | 특징 벡터 |

Abstract

This paper proposes visual-feature extraction for each band in wavelet domain with both spatial frequency features and multi resolution features, and the combination of visual features using fuzzy integral. In addition, it uses color feature expression method taking advantage of the frequency of the same color after color quantization for reducing quantization error, a disadvantage of the existing color histogram intersection method. Also, it is found that the final similarity can be represented in a linear combination of the respective factors(Homogram, color, energy) when each factor is independent one another. With respect to the combination patterns the fuzzy measurement is defined and the fuzzy integral is taken. Experiments are performed on a database containing 1,000 color images. The proposed method gives better performance than the conventional method in both objective and subjective performance evaluation.

■ keyword : | CBIR | Fuzzy Integral | Feature Extraction | Visual Features |

1. Introduction

Content-based Image Retrieval(CBIR) has been widely used since the 1992 proposal of Kato[1]. In CBIR, user can retrieve, easily, images that he or

she wants, by taking advantage of feature value representing visual properties like image color, shape, texture, etc. Therefore, the most important is effective extraction of such low-level visual features representing image as image color, shape, texture,

* 본 논문은 2006년도 교육인적자원부 지방연구중심대학 육성사업의 지원에 의하여 연구되었습니다.

접수번호 : #060323-001

접수일자 : 2006년 03월 23일

심사완료일 : 2006년 04월 12일

교신저자 : 김동우, e-mail : dubssi@paran.com

etc.

In 1991, Swain and Ballard[2] proposed a color histogram method, a typical method extracting features by using color information. Color histogram is easy to calculate, and robust even about slight camera location change. However, its disadvantage is that images with different spatial color distribution can have the same color histogram. CCV(Color Coherence Vector)[3] techniques tried to compensate the disadvantage, but a lot of calculation takes too much time.

Texture is a visual pattern of similar properties, not of an only property like color or brightness, and a property that all substantial surfaces have. In the early 1970s, Haralick[4] brought forward co-occurrence matrix as texture feature. The method took advantage of gray level as the spatial dependence of texture. First of all, it constructs a co-occurrence matrix based on distance and direction between image pixels. Then it extracts meaningful statistics from matrix for expressing texture.

Jacobs, etc.[5] took advantage of wavelet transformation, for the first time, in extracting feature vector. They made feature vector for each image channel R, G, and B by using Harr wavelet transformation fast to calculate. But this method has long trans-action time because of too big feature vector dimensions. To overcome, Park, etc[6] proposed a similarity measurement method using fuzzy theory, new color feature expression method, and visual- feature extraction method from wavelet domain, a transformation domain, for each band. And CBIR is embodied by extracted visual features. But, this method uses the manually fixed relative weight value of visual features(Homogram, color, energy).

Accordingly, this study has proposed CBIR

method using fuzzy integral based on the conventional visual features. The combination of 3 visual features is decided by fuzzy measure and fuzzy integral. And this approach shows stable and reliable retrieval results for CBIR.

Chapter 2 describes the proposed visual feature extraction method, and similarity measurement method ; Chapter 3 describes Simulation measurement; Chapter 4 shows a simulation result, and finally, Chapter 5 comes to a conclusion.

II. Visual-feature extraction method

1. Homogram & Energy features

The existing histogram method considered only the gray level value or color level value of each pixel within image, not spatial correlation between pixels; therefore, even images with different spatial pixel distribution can have the same histogram.

In order to overcome the disadvantage of the existing histogram method, Cheng, etc.[7] applied pixel-based fuzzy homogeneity function to image segmentation. It extracts fuzzy homogeneity function using the frequency component of LL band, wavelet transformation domain, and apply it to image retrieval.

It has transformed wavelet by using 9/7 tap biorthogonal wavelet filter[8] ; then got LL, LH, HL, and HH band. N-number feature vectors have been extracted by LL band feature able to take advantage of sufficient spatial correlation in the form closely resembling an original image. A transaction process before extracting N-number features is as following.

- step 1. Input image is normalized to $M \times N$ size.
- step 2. The normalized image is transformed into 1-level wavelet.

step 3. LL-band wavelet coefficients are normalized to L-level values.

If LL-band wavelet coefficients normalized to L-level values are defined as $c(i, j)$, it will result in the following eq.(1). Here, i and j mean wavelet coefficient index.

$$c(i, j) \in \{0, 1, 2, 3, \dots, L-1\}, \quad 0 \leq i \leq M/2, 0 \leq j \leq N/2 \quad (1)$$

Using normalized LL-band wavelet coefficients, L-number homogram values are obtained in the LL band of wavelet domain; then stored in image database.

Also, the study extracted 3 feature vectors by high-frequency band in wavelet band. As energy is an instrument measuring the uniformity of brightness, the energy value of each band is obtained and stored in image database by its relevant image feature vector. Energy value is obtained by the square value of each coefficient. Therefore, the energy value of each wavelet band is obtained by eq.(2) - here $B \in \{LH, HL, HH\}$.

$$E(B) = \sum_{(i,j) \in E} c^2(i, j) \quad (2)$$

2. Color feature extraction techniques

The existing color histogram takes advantage of only the frequency of the same color, not color component like relevant color R, G, or B value, or hue value. As a result, organizing color histogram gets to absorb quantization error, as it is, resulting from color quantization. Accordingly, Park, etc. proposed a method using every color value and frequency in order to overcome this disadvantage.

It took advantage of HSV color space able to

divide brightness and color difference for making color feature more obvious; uniformly quantized H, S, V value respectively into 4, 2, and 2 values, and then divided into 16 color regions for dividing the image region including similar colors. The representative color and frequency of each divided region were extracted as its features, and each representative color had the average value R, G, and B of each region.

Color feature extracted from image is expressed in the type of vector as the following eq.(3).

$$\text{color feature} = [r_i, g_i, b_i, p_i], \quad i=1, 2, \dots, M \quad (3)$$

Here, M means the total number of color; p_i means the ratio value of a relevant color with r , g , or b ; they all satisfy $\sum p_i = 1$. The size of color feature vector extracted from one image is $M \times 4$ byte, occupying a tiny space.

III. Simulation measurement

1. General method

The purpose of CBIR system is retrieving the most similar image to query image within image database. Therefore, the method of obtaining similarity between query image and target image has a big influence on the performance of the total system. The study has proposed distance calculation (similarity calculation) method fit for 3 feature components like color, homogram, and energy value.

As the color distance calculation method using the existing color histogram takes advantage of only the frequency of the same color, not color components like relevant color values(R, G, or B), even similar

visual images no longer have similarity because of color histogram movement by such component as illumination change.

Color histogram applies both the RGB value and frequency of each color to fuzzy color function, and measures similarity between query image and target image. If query image and target image are respectively I_1 and I_2 , each image color feature can be expressed as [r, g, b, p]. At this time, similarity between the colors of two images is obtained by membership function like eq.(4). If two colors are similar, eq.(4) has a value of about 1 ; if not, it has 0.

$$\mu_A(c_1, c_2) = e^{-\{(r_1-r_2)^2+(g_1-g_2)^2+(b_1-b_2)^2\}} \quad (4)$$

Eq.(5), distance function D_2 transformed by similarity between these two color values, can find similarity between two images.

$$D^2(I_1, I_2) = \sum_{i=1}^B p_{1i}^2 + \sum_{j=1}^B p_{2j}^2 - \sum_{i=1}^B \sum_{j=1}^B 2\mu(I_1, I_2)p_{1i}p_{2j} \quad (5)$$

Here, I_1 is the color feature of query image, I_2 the color feature of target image, and B the number of colors.

The homogram set, H, of one image has 128 bins like in H(0), H(1), ..., H(127) ; if 128 bins are all compared at the time of calculating homogram distance between two images, calculation gets to increase much more. Therefore, this study calculated homogram distance between two images by using the biggest 32 homogram values after arranging query image homograms according to frequency, not calculating similarity by the homogram set of query image. This is expressed in eq.(6).

$$D_H(Q, I) = \sum_{t \in T_Q} |H_Q(t) - H_I(t)| \quad (6)$$

Here, Q means query image, and I target image. T_Q means 32 homogram indexes with the most biggest frequency of the homogram values of image Q.

Energy distance between query image and target image was obtained by eq.(7) using general Euclidian distance.

$$D_E(Q, I) = \sum_{B \in \{LH, HL, HH\}} |E_Q(B) - E_I(B)| \quad (7)$$

The total distance between two images is obtained by the total difference in the energy value of each band of wavelet-transformed image.

2. The proposed method

The purpose of CBIR system is retrieving the most similar image to query image within image database. Therefore, the method of obtaining similarity between query image and target image has a big influence on the performance of the total system. This study proposes distance calculation (similarity calculation) method fit for 3 feature components like color, homogram, and energy. We use Sugeno's[9] fuzzy integral and fuzzy measure in order to combine three features.

Since fuzzy measure and fuzzy integral, generalizing the usual definition of a classical measure by replacing the usual additive property by a weaker requirement, i.e the monotonicity property with respect to set inclusion, the concept of them has been often used in decision making. Subjective evaluation models using fuzzy integrals have been applied in various fields, and their effectiveness has been experimentally shown.

Fuzzy integral and fuzzy measure is explained as follows. In this paper, min and max are denoted \wedge , \vee respectively. Let X be a non-empty set, \mathcal{A} be a σ -algebra of subsets of X .

A set function $g : \mathcal{A} \rightarrow [0, 1]$ is called fuzzy measure if

- (1) $g(\Phi) = 0$
- (2) $A, B \in \mathcal{A}$ with $A \subset B$ implies $g(A) \leq g(B)$.

We will assume here $g(X) = 1$ as usual, although this is not necessary in general.

We call (X, \mathcal{A}, g) a fuzzy measure space if g is a fuzzy measure on a measurable space (X, \mathcal{A})

Note that the usual additivity axiom for probability measures

$$g(A \cup B) = g(A) + g(B), \quad A \cap B = \Phi \quad (8)$$

has been replaced by a weaker one: *monotonicity*. For this reason, the definition of a fuzzy measure requires in general 2^n coefficients, namely the measure of the 2^n subsets of X . Fuzzy measures include as particular cases probability measures, possibility and necessity measures, belief and plausibility functions, etc.

A convenient way of representing fuzzy measures in the finite case is to use a lattice representation. All the 2^n coefficients defining a fuzzy measure can be arranged in a lattice with the usual ordering on real numbers, which is in fact the same as the Boolean lattice of subsets of X , ordered with inclusion.

A real-valued function $h: X \rightarrow [0, 1]$ is \mathcal{A} measurable with respect to \mathcal{A} and \mathcal{B} (measurable,

for short, if there is no confusion likely) if

$$h^{-1}(B) = \{x | h(x) \in B\} \in \mathcal{A} \quad \text{for any } B \in \mathcal{B} \quad (9)$$

where \mathcal{B} is the σ -algebra of Borel subsets of $[0, 1]$.

We turn now to fuzzy integrals, that is, integrals of a real function with respect to a fuzzy measure, by analogy with Lebesgue integral which is defined with respect to an ordinary(i.e. additive) measure. Sugeno has proposed the following definition, restricted to functions on $[0, 1]$:

$$L^0(X) = \{ h : X \rightarrow [0, 1] \mid h \text{ is measurable with respect to } \mathcal{A} \text{ and } \mathcal{B} \} \quad (10)$$

For $h \in L^0(X)$ the fuzzy integral of on A with respect to a fuzzy measure g defined by

$$\int_A h dg = \sup_{\alpha \in [0,1]} [\alpha \wedge g(A \cap H_\alpha)] \quad (11)$$

When $A=X$, the fuzzy integral is denoted by $\int h dg$.

We present the algorithm for evaluation using the fuzzy integral which is explained above.

step1. an object be given to evaluate, a factor space of the object $X = x_1, x_2, \dots, x_n$,

In this paper, let $n=3$, x_1 is homogram, x_2 is color, and x_3 is energy.

step 2. given an importance measure as follows: for each subset A of the factor space X , a real number $0 \leq g(A) \leq 1$ be given as a degree of importance of A , such a real number should be the maximum possible score that the object can gain relying only on the quality factors in A .

step 3. for each individual quality factor

x_1, x_2, \dots, x_n we enumerate evaluated scores in the order $h(x_1), h(x_2), \dots, h(x_n)$ as following :

For $x_i \in X, i = 1, 2, \dots, n$, let

$$h(x_1) \leq h(x_2) \leq \dots \leq h(x_n)$$

then we can consider the set

$$H_i = \{x_k | k = i, i+1, i+2, \dots, n.\}$$

step 4. we compute $h(x_i) \wedge g(H_i)$ for each i .

step 5. we can take the maximum value of all computed values in step 4 that is the fuzzy integral $\int h dg$. of the scores $h(x_i)$ with respect to the importance measure g to obtain a synthetic evaluation of the quality of the given object.

IV. Simulation results

[Fig. 1] shows block diagram of the proposed method. Above all, one input image is normalized to $M \times N$ size ; the normalized image is transformed into HSV color space for extracting color feature, then the image is divided into 16 color regions, and the representative color of each region is calculated; color histogram is obtained by the above result. Color histogram is expressed in the form of vector, and stored in image feature vector database.

Also, color image is transformed into black and white image for calculating homogram and energy.

Black and white image is decomposed into 1-level by wavelet transformation; then feature vector fit for each band is extracted. First of all, LL band extracts 128 homogram vectors using spatial correlation between wavelet coefficients in the form closely resembling an original image; The other high-frequency band LH, HL, HH, etc. extract 3 feature vectors by energy value, and store them in

image feature vector database.

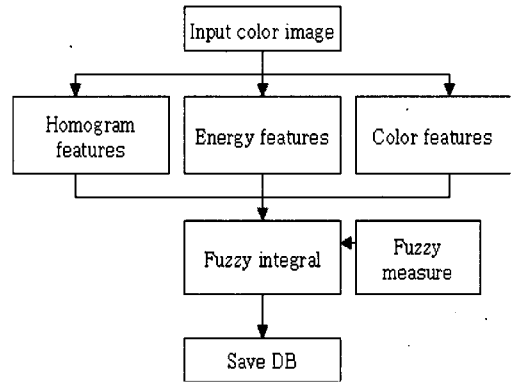


Fig. 1. Block diagram of the proposed method.

This study has used 1,000 natural images[10] such as buses, horses, flowers, characters with African scenes for background, European townscape, mountains, the sea, the seaside, etc. for evaluating the performance of content-based image retrieval system taking advantage of the feature extraction techniques and similarity measurement techniques proposed here.

Also, it used a system absorbing both edge histogram method expressing high-frequency component and the existing Swain's histogram intersection method for evaluating the objective performance of CBIR system using the proposed feature extraction techniques. The existing system experimented with color histogram weight and edge histogram weight respectively 0.6 and 0.4, while the method proposed here has taken advantage of query transaction techniques by user definition, and experimented with color, homogram, and energy weight each as 0.4, 0.4, and 0.2.

This paper compared the technique of extracting color-feature and the performance of retrieving similarity proposed here with those of Swain's histogram intersection method and Park's visual

feature method, and more.

Table 1. Comparison of the retrieval results

Images	Color histogram		Park's method		Proposed method	
	Precision	Recall	Precision	Recall	Precision	Recall
Bus	0.5	0.25	0.9	0.45	0.92	0.47
Horse	0.78	0.38	0.89	0.4	0.91	0.43
Flower	0.87	0.42	0.91	0.47	0.96	0.48

[Table 1] shows the result of having retrieved, of 1,000 experiment images, ones similar to 10 query images voluntarily selected from bus-kind, and horse-kind, flower-kind ones. It has been proved that the proposed method has better recall and precision than the existing.

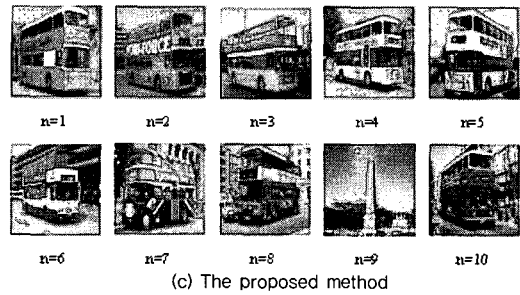
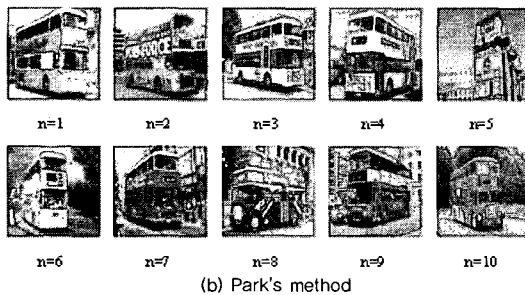
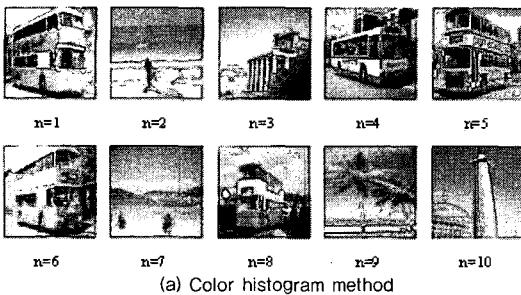


Fig. 2. Simulation result(bus)

Also, when [Fig. 2] has received bus image as query image for subjective estimation, (a), (b) and (c) respectively show the result of retrieving the color histogram method, Park's method and proposed method. As a result of retrieval, images have been expressed as index according to their similarity.

That is, the lower the index of an image is, the more similar it is to query image. The color histogram method retrieved 5 images, like #1, #4, #5, #6, #8, etc. similar to query image (n=1), of the 10 retrieval images. And Park's method retrieved 9 similar images. On the other hand, the proposed method has showed better visual result than Park's method as it retrieved 9 similar images.

V. Conclusion

The existing CBIR has used color, texture, and shape etc. Most of all used each features, independently. But the proposed method is used all of the 3 visual features(color, homogram, energy). And it is named as the similarity retrieval method using fuzzy integral, which can be represented in a linear combination of the 3 visual features.

The study experimented 1,000 color images by proposed method here; as a result, the new methods have proved to be much more precise in both

objective performance evaluation and subjective performance evaluation method than the existing ones.

Reference

[1] K. Hirata and T. Kato, "Query by visual example," Proceedings of EDBT Conference, pp.56-71, 1992.

[2] M. J. Swain and D. H. Ballard, "Color indexing," International Journal of Computer Vision, Vol.7, No.1, pp.11-32, 1991.

[3] G. Pass, R. Zabih, and J. Miler, "Comparing images using color coherence vectors," Proceedings of ACM Conference on Multimedia, pp.65-73, 1996.

[4] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Texture features for image classification," IEEE Trans. on Systems, Man, and Cybernetics, Vol.3, No.6, pp.610-621, Nov., 1973.

[5] C. E. Jacobs, A. Finkelstein, and D. H. Salesin, "Fast multiresolution image querying," In Proceeding of SIGGRAPH 95, pp.276-286, Aug., 1995.

[6] W. B. Park, "Visual feature extraction technique for content-based image retrieval," Chungbuk National University thesis, 2004.

[7] H. D. Cheng and C. H. Chen, "Fuzzy homogeneity approach to multilevel thresholding," IEEE Trans. Image Processing, Vol.7, No.7, pp.1084-1086, July, 1998.

[8] J. D. Villasenor, B. Bellzer, and J. Liao, "Wavelet filter evaluation for image compression," IEEE Trans. on Image Processing, Vol.4, No.8, pp.1053-1060, Aug., 1995.

[9] M. Sugeno, "Fuzzy Measures and Fuzzy Integrals: A Survey," in Fuzzy Automata and Decision Processes, pp.89-102, 1977.

[10] J. Z. Wang, J. Li, and G. Wiederhold, "SIMPLiCity: Semantics-Sensitive Integrated Matching for Picture Libraries," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.23, No.9, pp.947-963, 2001.

저 자 소 개

송 영 준 (Young-Jun Song)

중신회원



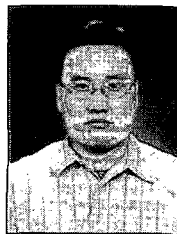
- 1994년 2월 : 충북대학교 정보통신공학과(공학사)
- 1996년 8월 : 충북대학교 정보통신공학과(공학석사)
- 2004년 8월 : 충북대학교 정보통신공학과(공학박사)

• 2005년 9월~현재 : 충북대학교 전기전자컴퓨터공학부 박사 후 연구원

<관심분야> : 영상인식, 영상처리, 얼굴인식

김 동 우 (Dong-Woo Kim)

정회원



- 1997년 2월 : 충북대학교 정보통신공학과(공학사)
- 2002년 2월 : 충북대학교 정보통신공학과(공학석사)
- 2006년 2월 : 충북대학교 정보통신공학과(공학박사)

• 2006년 3월~현재 : 충주대학교 시간강사

<관심분야> : 내용기반검색, 멀티미디어 정보처리

김 남(Nam Kim)

정회원



- 1981년 : 연세대학교 전자공학과 (공학사)
 - 1983년 : 연세대학교 전자공학과 (공학석사)
 - 1998년 : 연세대학교 전자공학과 (공학박사)
 - 1992년~1993년 : 미 Stanford 대학 방문교수
 - 2001년~2002년 : 미 California Technology Institute 방문교수
 - 1989년~현재 : 충북대학교 전기전자컴퓨터공학부 교수
- <관심분야> : 디지털 이동통신, 이동통신, 무선 시스템, 전자파해석, EMI/EMC, Diffractive diffractive optics, and optical memory system

김 미 혜(Mi-Hye Kim)

종신회원



- 1992년 2월 : 충북대학교 수학과 (이학사)
 - 1994년 2월 : 충북대학교 수학과 (이학석사)
 - 2001년 2월 : 충북대학교 수학과 (이학박사)
 - 2001년 4월~2004년 8월 : 충북대학교 전기전자컴퓨터공학부 초빙교수
 - 2004년 9월~현재 : 충북대학교 전기전자컴퓨터공학부 조교수
- <관심분야> : 퍼지이론, 정보이론, 금융수학