

論文2000-37SP-4-4

## 로젯 패턴을 이용한 영상 검색 기법

## (Image Retrieval Using the Rosette Pattern)

姜應寬\*, 張聖甲\*, 宋浩根\*\*, 崔宗秀\*

(Eung Kwan Kang, Surng Gabb Jahng, Ho Keun Song,  
and Jong Soo Choi)

## 요약

본 논문에서는 영상에 대해 영상 중심 부분에 주사가 집중되는 특성을 지닌 로젯 패턴을 적용함으로써 영상의 전체적인 그레이 분포 특징뿐만 아니라 국부적인 위치 정보까지 얻을 수 있으며 이를 히스토그램 분포 특성으로 변환시켜 데이터베이스의 색인 정보로 저장한 후, 입력되는 질의 영상에 대해 사용자가 원하는 검색 결과를 제시하는 방법에 대해 제안한다. 실험 결과 제안된 방법이 검색에 있어 우수한 성능을 갖추고 또한 색인에 필요한 메모리 공간을 크게 줄일 수 있음을 확인하였다.

## Abstract

This paper presents a new indexing technique, for the fast content-based image browsing and retrieval in a database. By applying the rosette pattern that has more sample lines in the vicinity of center than those in the outer parts, we can get global gray distribution features as well as local positional information. These features are transformed into histogram and used as database indices. From the simulation results, the proposed method clearly shows the validity and the efficiency in respect of memory space as well as a good retrieval performance.

## I. 서론

최근 영상 압축 및 통신 기술이 급속히 발전함에 따라 인터넷, VOD, 디지털 전자도서관 등 여러 분야에

걸쳐 디지털 영상 정보의 활용이 폭발적으로 증가하고 있으며, 이를 효과적으로 관리하고 검색하기 위한 방법이 활발히 연구되고 있다<sup>[1]</sup>. 또한, 효율적인 데이터의 관리 및 검색 등을 위해 정보의 특징 내용을 갖고 대상을 표현하고자 하는 MPEG-7 표준 작업도 현재 진행 중에 있다<sup>[2]</sup>.

\* 正會員, 中央大學校 尖端映像大學院 映像工學科  
(Dept. of Image Engineering, Graduate school of Advanced Imaging Science, Multimedia and Film, Chung-Ang University)

\*\* 正會員, 韓瑞大學校 컴퓨터情報學科  
(Dept. of Computer Science and Engineering, Hanseo University)

※ 본 연구는 교육부 BK21사업의 일부 지원 및 정보통신부 정보통신 우수시범학교 지원사업의 일환으로 수행된 결과입니다.

接受日字:2000年1月6日, 수정완료일:2000年5月22日.

한편, 영상의 특징 정보를 추출하기 위해서는 먼저 입력 영상을 받아들여 읽어들이는 즉, 주사 (scanning) 하는 과정이 선행되어야 한다. 기존의 주사 방법으로 순차 주사 패턴 (progressive scan pattern)이나 비월 주사 패턴 (interlaced scan pattern), 혹은 압축에 주로 이용되는 지그재그 주사 패턴 (zigzag scan pattern)<sup>[3]</sup> 등에 있는데 이들 모두는 영상을 구성하는 전체 픽셀 정보를 읽어들이는 방식이다. 이에 반해 본 논문에서는 인간 시각 특성상 영상의 외곽 부분보다는 중심 부분에서 보다 신경이 민감한 점에 착안하여, 영상 중심 부분에 주사가 집중되는 특성을 지닌 로젯 주사 패턴

(rosette scan pattern)을 이용하여 영상의 내용을 읽어 들이고 이로부터 특징 정보를 추출하여 데이터베이스의 색인 정보로 저장한 후, 입력되는 질의 (query) 영상에 대해 사용자가 원하는 유사 영상을 제시하는 영상 검색 기법에 대해 제안한다.

본 논문의 구성은 다음과 같다. 먼저 2장에서 로젯 주사 방식에 대한 이론적 배경과 실제 컴퓨터 상에서 로젯 패턴을 생성하는 방법에 대해 기술하고, 3장에서는 로젯 패턴을 이용한 영상 색인 및 검색 기법에 대해 설명한 후 로젯 패턴 적용에 대한 타당성 및 효용성에 대해 살펴본다. 4장에서는 제안한 방식의 이용한 성능 평가 및 분석 내용을 검토하고, 5장에서 결론 및 향후 연구 방향에 대해 언급한다.

## II. 로젯 주사 방식의 이론적 배경

로젯 주사 방식은 레이더, 적외선 신호 처리 및 의료 진단 장치에 주로 사용되는 기법으로, 로젯 패턴은 두 개의 광학계를 일정한 비율로 서로 반대 방향으로 회전시킴으로써 이루어진다<sup>[4][5]</sup>. 이에 근간하여 컴퓨터 상에서 구현할 경우 즉, 회전 주파수를  $f_1$  과  $f_2$  로 회전시킬 때 형성되는 로젯 패턴의 자취 방정식은 다음과 같이 극 좌표계로 표현할 수 있다.

$$\begin{aligned} r(t) &= \delta \cos \pi(f_1 + f_2)t \\ \theta(t) &= \pi(f_1 - f_2)t \end{aligned} \quad (1)$$

이러한 회전 주파수 비에 의해 주사 속도, 총 잎사귀의 수, 잎사귀의 넓이 및 인접한 잎사귀와의 교차점들의 위치 등이 결정되며, 이들 주파수 비의 관계는 다음과 같이 나타낼 수 있다.

$$\frac{N_2}{N_1} = \frac{f_2}{f_1} \quad (2)$$

여기서  $N_1$  및  $N_2$  는 양의 정수이고  $f_1 > f_2$  이다.

식(2)의 비가 항상 유리수이고,  $N_1 = f_1/f$  및  $N_2 = f_2/f$  가 되는 최대 공약수  $f$ 를 가질 때에만 식(1)에 의해 완전히 닫혀 있는 로젯 패턴을 형성할 수 있다. 이 경우 로젯 패턴의 프레임 주기는

$$T = \frac{1}{f} = \frac{N_1}{f_1} = \frac{N_2}{f_2} \quad (3)$$

이고, 로젯 패턴 내의 전체 잎사귀의 수는

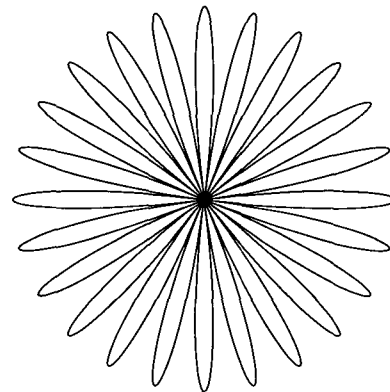
$$N = N_1 + N_2 \quad (4)$$

이 된다. 또한 잎사귀의 넓이를 나타내는 파라미터는 다음과 같이 나타낼 수 있다.

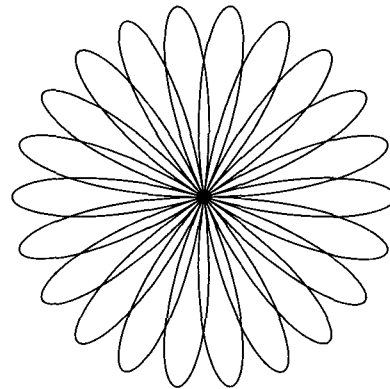
$$\Delta N = N_1 - N_2 \quad (5)$$

여기서  $\Delta N$ 은 잎사귀의 절대적인 넓이가 아닌 상대적인 척도이다.

그림 1은 위 수식에 근거하여  $N_1$ ,  $N_2$  및  $\Delta N$ 에 따라 생성된 로젯 패턴의 모양을 보여주고 있다.  $\Delta N$ 의 값이 커지면 잎사귀의 폭이 넓어지고, 값이 적어지면 상대적으로 잎사귀의 폭은 좁아진다.



(a)  $N_1 = 13, N_2 = 11, \Delta N = 2$



(b)  $N_1 = 13, N_2 = 9, \Delta N = 4$

그림 1. 로젯 패턴의 예  
Fig. 1. Examples of generated rosette pattern.

그림 1에서 알 수 있듯이 로젯 패턴은 패턴의 중앙 부분으로 갈수록 외곽에 비해 더 많은 주사 수 (scan

point)를 포함하는 특징을 갖고 있다. 이러한 로젯 패턴의 특징을 이용하여 영상 색인 및 검색에 적용시키는 방법에 대해 다음 장에서 기술하고자 한다.

### Ⅲ. 로젯 패턴을 이용한 영상 색인 및 검색 기법

일반적인 영상의 경우, 영상 취득 단계에서부터 이미 인간 인지시각 체계 (human psycho-visual system)에 근간하여 영상내 주요 객체 (object)는 중심부에 놓이도록 한다는 가정으로부터 영상의 주요 정보는 중심부에 집중한다는 사실을 유추할 수 있다<sup>[6]</sup>. 이러한 사실에 비추어 볼 때 로젯 패턴을 영상에 적용할 경우 정보의 기술 (description) 측면에서 기존 주사 패턴 방식에 비해 높은 효율성과 타당성을 갖는다는 것은 극명하다. 즉, 영상에 대한 픽셀 정보가 한정되어 있을 경우 기존의 순차 주사 패턴 (progressive scan pattern)이나 비월 주사 패턴 (interlaced scan pattern), 혹은 압축에 주로 이용되는 지그재그 주사 패턴 (zigzag scan pattern) 등에 비해 로젯 주사 패턴은 중심부에 많은 픽셀 정보를 할당하므로 보다 정확한 정보의 기술이 가능하다. 따라서 영상에 대해 픽셀 전체를 스캔하지 않고도 주요 정보를 포함시킬 수 있으므로, 영상의 특징 정보를 색인 (indexing)할 경우 필요한 메모리 양을 기존 주사 패턴에 비해 크게 줄일 수 있음을 의미한다.

로젯 패턴을 영상 색인 및 검색에 적용하기 위해 제안하는 방법의 기본 개념도를 그림 2에 나타내었다.

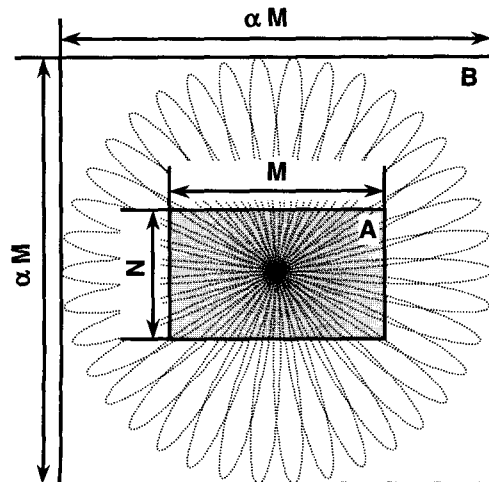


그림 2. 개념도  
Fig. 2. Concept diagram.

그림 2에서  $M, N$  은 대상 영상의 크기를 나타낸다. 여기서,  $\alpha$  는 크기 인자 (scale factor)를 나타내는데 이는 로젯 패턴의 특성상 가로·세로 비가 정방형의 비율을 갖는데 반해, 일반적인 영상, 특히 TV 영상인 경우 종횡 비 (aspect ratio)가 4:3 혹은 16:9 정도의 비율을 가지므로 대상 영상을 로젯 패턴의 중심부에 놓이도록 하기 위함이다. 또한, B는 전체 로젯 패턴이 주사되는 영역을 나타내며, A는 B영역 가운데 일부분 즉, 대상 영상 크기에 맞도록 패턴이 마스크된 영역을 나타낸다. 제안하는 로젯 패턴을 이용한 영상 색인 방법에 대해 보다 구체적인 내용을 기술하면 다음과 같다. 먼저 로젯 패턴의 자취 방정식으로부터 2차원 직교 좌표계에서의 패턴의 궤적을 식(6)을 이용하여 구한다.

$$\begin{aligned} x(t) &= \frac{\delta}{2} (\cos \omega_1 t_s + \cos \omega_2 t_s) \\ y(t) &= \frac{\delta}{2} (\sin \omega_1 t_s - \sin \omega_2 t_s) \end{aligned} \quad (6)$$

여기서,  $\omega_1$  과  $\omega_2$  는 패턴을 형성하는 회전 속도를 나타내며,  $t_s$  는 연속 샘플링 간격 (continuous sampling time)을 나타낸다.  $M \geq N$  인 경우  $\delta = \alpha M$ ,  $M < N$  인 경우  $\delta = \alpha N$  으로 하였으며, 이때의  $M, N$  은 대상 영상의 크기를 나타낸다. 이로부터 계수화된 로젯 패턴 영상 (digitized rosette pattern image), 즉 그림 2에서의 B영역에 해당하는 주사점 (scan point)에 대한 배열 순서 (array number)는 다음의 식(7)과 같이 나타낼 수 있다.

$$\begin{aligned} I &= \lfloor x(t_s) + \delta \rfloor \\ J &= \lfloor y(t_s) + \delta \rfloor \end{aligned} \quad (7)$$

여기서  $\lfloor Z \rfloor$  는  $Z$ 를 넘지 않는 최대 정수를 의미한다. 다음으로 전체 로젯 패턴 영상으로부터 대상 영상 크기에 맞도록 패턴 마스크된 영역, 즉 그림 2에서의 A영역만을 취하기 위해,  $I \leq M$  과  $J \leq N$  의 조건을 만족하는 픽셀에 대해서만 식(8)과 같이 메모리에 저장된다.

$$S(n) = \text{pixel}(I, J) \quad \text{if } I \leq M, J \leq N \quad (8)$$

여기서  $n$  은 0 이상의 정수이다. 그림 3은 원영상에 대해 제안하는 로젯 패턴을 적용시킨 영상과 이들 각각에 대한 히스토그램 분포를 나타낸 것이다.

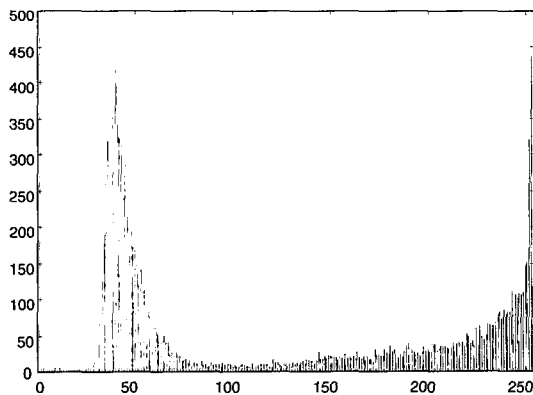
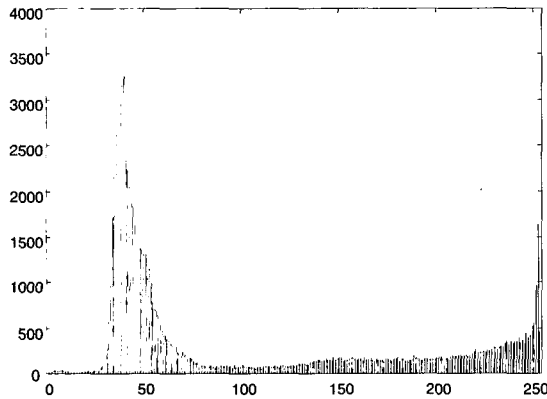
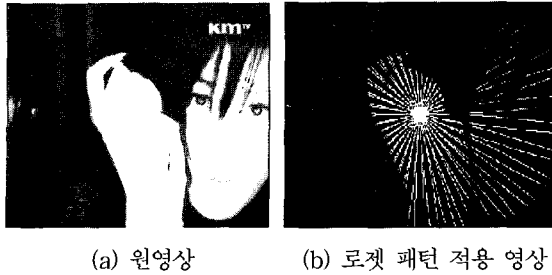


그림 3. 원영상과 로켓 패턴을 적용한 영상의 예  
Fig. 3. Example of original and rosette pattern masked image.

한편, 그림 3의 (c)와 (d)에서 볼 수 있듯이 로켓 패턴을 적용한 영상의 그레이 (Gray) 히스토그램 분포 특성은 원영상에 대한 전체 히스토그램 분포 특성과 거의 일치함을 알 수 있다. 이는 로켓 패턴의 특성상 주사되는 픽셀 수 (scanning pixel point)가 영상의 외곽 부분뿐만 아니라 중심부로 갈수록 더욱 집중되기 때문에, 입력 영상에 대해 픽셀 전체를 주사하지 않고

제안하는 방법에 따라 로켓의 궤적만 취하더라도 영상에 대한 전역적인 정보 (global information)를 충분히 포함시킬 수 있는 것으로 해석할 수 있다. 또한 로켓 패턴의 궤적이 지나가는 자취점은 고유한 것으로 이는 영상내 객체들의 국부적인 위치 정보 (local positional information)를 포함하고 있다는 사실도 동시에 알 수 있다. 더욱이 영상 색인 (indexing)에 사용되는 픽셀의 개수는 기존 주사 방식에 의한 것보다 크게 줄일 수 있으므로 메모리 측면에서도 매우 효율적임을 알 수 있다. 따라서 본 논문에서는 영상에 대해 로켓 패턴을 적용시킨 후 이에 대한 그레이 분포 특징을 히스토그램으로 변환시켜 데이터베이스의 색인 정보로 저장한다. 다음으로 임의의 질의 영상에 대해 로켓 패턴을 적용시켜 히스토그램 분포를 구하고, 데이터베이스에 저장된 각각의 히스토그램 분포와의 유클리드 거리 차를 아래의 식(9)를 이용하여 계산한 후 값이 작을수록 유사도가 높은 것으로 순위를 매김으로써 검색을 수행하였다.

$$d(Q, I) = \left[ \sum_{k=1}^n (Q_k - I_k)^2 \right]^{1/2} \quad (9)$$

여기서,  $Q_k$ 는 질의영상의 히스토그램을,  $I_k$ 는 데이터베이스에 저장된 히스토그램을 각각 나타내며,  $k$ 는 히스토그램 빈 (bin)을 나타낸다.

#### IV. 실험 및 고찰

본 장에서는 논문에서 제안된 로켓 패턴을 이용한 영상 검색 기법의 성능을 평가하기 위해 320×240 크기를 갖는 1240프레임의 뮤직 비디오 영상과 전체 7421 프레임의 뮤직 비디오 영상에 대해 [7]에 의해 추출된 115프레임의 대표프레임에 대해 256 그레이레벨 해상도를 갖는 영상을 대상으로 실험한 결과에 대해 기술한다.

표 1은 로켓 패턴의 일사귀 수에 따른, 색인에 사용된 픽셀 수와 평균 검색율 (average retrieval ratio, e.g. precision)을 비교한 것이다. 여기서의 평균 검색율이란 다음의 식(10)과 같이 정의된다<sup>[8]</sup>.

$$P_{\%} = \frac{R_n}{T_n} \times 100 \quad (10)$$

여기서  $T_n$ 은 데이터베이스에서 검색된 영상의 총 개

수,  $R_n$  은 정확히 검색된 영상의 개수를 각각 나타낸다. 먼저, 정확한 검색 여부를 객관적으로 판단하기 위해 비디오에서 장면을 구성하는 최소 단위인 샷 (shot)에 대해 우선적으로 검색을 수행하였다. 일반적으로 뮤직비디오인 경우 하나의 샷은 보통 60~70프레임 정도로 구성이 되며, 동일 샷내의 프레임들은 주어진 일정 구간에서 서로 비슷한 공간적 동질성을 지닌다. 따라서 하나의 샷내의 임의의 영상이 질의로 주어질 경우, 주어진 질의 영상이 속한 샷의 영상들을 모두 찾아내는지를 확인해야 한다. 물론 뮤직 비디오의 특성상 하나의 장면과 유사한 장면이 또다른 구간, 즉 다른 샷에서 반복될 수도 있으나 이 경우는 정확히 검색된 영상의 개수에 포함을 시키지 않았다. 한편, 동일 실험 기준, 즉 임의의 질의 영상에 대해 주어진 질의 영상이 포함된 샷의 영상들을 기존의 일반적인 히스토그램, 즉 영상 전체에 대한 분포특성을 이용하여 식(9)를 계산한 후 검색할 경우 약 73.7%의 평균 검색율을 보였다.

또한, 크기 인자 (scale factor)에 따른, 색인에 사용된 픽셀 수와 평균 검색율을 비교하여 표 2에 나타내었다.

표 1. 로젯 잎사귀 수에 따른 색인 픽셀 수 및 평균 검색율 ( $\alpha = 2$ 일 때)

Table 1. The number of indexed pixels and precision according to the number of rosette leaves (when  $\alpha = 2$ ).

비교 항목 \ 로젯 패턴 파라미터	( $M = 13$ $N_2 = 9$ )	( $M = 21$ $N_2 = 17$ )	( $M = 61$ $N_2 = 57$ )
색인 픽셀 수	7,614	12,298	26,797
평균 검색율(%)	96.0	94.7	97.3

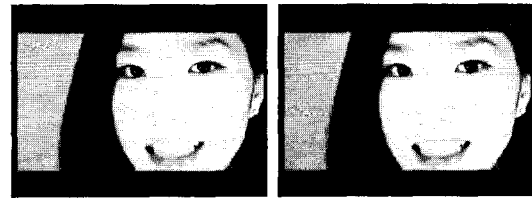
표 2. 크기 인자에 따른 픽셀 수 및 평균 검색율 ( $N_1=13, N_2=9$ 일 때)

Table 2. The number of indexed pixels and precision according to the scale factor (when  $N_1=13, N_2=9$ ).

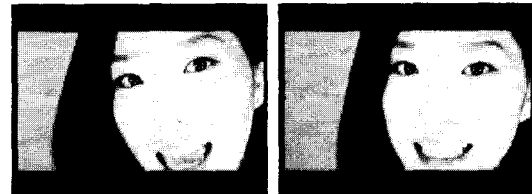
비교항목 \ 크기 인자 ( $\alpha$ )	1	1.5	2	3
색인 픽셀 수	7,745	7,866	7,614	7,397
평균 검색율(%)	96.0	96.0	96.0	94.7

따라서 실험에 사용한 로젯 패턴은 색인에 사용된 픽셀 수가 적으면서 동시에 평균 검색율은 높은 것을

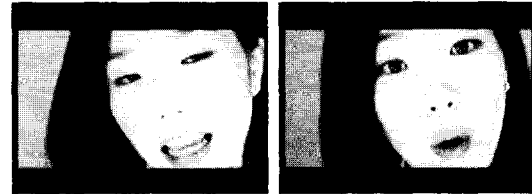
선택하였다. 즉, 표1과 표2에서 살펴보듯이 로젯 패턴의 전체 잎사귀 수는 22개 ( $N_1=13, N_2=9$ 인 경우)이며, 크기 인자는 2를 설정하였다. 한편, [7]에 의해 비디오에 있어 장면 전환이 발생하는 부분, 즉 샷 (shot)의 경계를 찾아내고 각 샷의 장면을 대표하는 대표프레임을 추출한다. 다음으로 추출된 대표프레임에 대해 제안하는 로젯 패턴을 적용한 후 히스토그램 분포 특성을 데이터베이스의 색인 정보로 저장한다. 따라서 사용자가 임의의 한 프레임을 질의 영상으로 선택하면 이에 대해 유사한 대표프레임을 식(9)의 계산을 통해 유사도가 높은 순서대로 검색을 수행한다. 다음의 그림 4는 임의로 주어진 질의 영상 (a)에 대해 [7]에 의해 추출된 대표 프레임들 가운데 가장 유사도가 높은 순위별로 검색된 결과를 나타낸 것이다.



(a) 임의의 질의 영상 (#2116) (b) 1순위 대표프레임 (#2117)



(c) 2순위 대표프레임 (#2106) (d) 3순위 대표프레임 (#2129)



(e) 4순위 대표프레임 (#2091) (f) 5순위 대표프레임 (#4513)



(g) 6순위 대표프레임 (#4754) (h) 7순위 대표프레임 (#3236)

그림 4. 검색 결과  
Fig. 4. Retrieval results.

로젯 패턴을 이용한 평균 검색율은 96.0%로 기존의 일반적인 히스토그램 분포를 이용한 방식에 비해 우수한 검색 성능을 지닌다. 더욱이 기존의 영상 전체에 대한 픽셀을 색인에 이용할 경우 76,800개의 픽셀이 필요하였으나 제안하는 방법에서는 약 10% 정도인 7614개의 픽셀만이 사용되었다. 그 밖의 실험을 통해서도 제안한 방법이 검색에 있어 매우 효과적이며, 특히 색인에 필요한 메모리 공간 측면에서도 아주 효율적임을 확인하였다.

## V. 결 론

본 논문에서는 로젯 패턴을 이용하여 영상의 그레이 특징 정보를 히스토그램 분포로 색인하고, 각각의 히스토그램 분포에 대한 거리 차를 구함으로써 유사 영상을 검색하는 방법에 대해 제시하였다. 제안한 방법은 검색에 있어 우수한 성능을 지니며, 또한 색인에 필요한 픽셀 수를 크게 줄임으로써 메모리 측면에서도 매우 효율적임을 확인하였다. 향후 제안한 방법은 색상(color)이나 질감(texture)과 같은 다른 색인 정보와 결합할 경우 보다 나은 영상 색인과 검색 수단으로 제공될 것이다.

## 참 고 문 헌

[1] A.K. Elmagarmid et al., *Video Database Systems: Issues, Products, and Applications*, Kluwer Academic Publishers, 1997.

[2] ISO/IEC JTC1/SC29/WG11 N2082 *MPEG-7 Context and Objectives*.  
 [3] ISO/IEC JTC1/SC29/WG11 13818-2.  
 [4] G.J. Zissis and W.L. Wolfe, *The Infrared Handbook*, Office of Naval Research Department of Navy, Arlington, VA., 1985.  
 [5] S.G. Jahng, H.K. Hong, S.H. Han, and J.S. Choi, "Design and analysis of improved instantaneous field of view of rosette scanning infrared seeker," *IEE Electronics Letters*, Vol. 33, No. 23, pp. 1964-1965, 1997.  
 [6] M. Stricker and A. Dimai, "Color indexing with weak spatial constraints," *Proc. SPIE*, Vol. 2670, pp. 29-39, 1996.  
 [7] E.K. Kang, S.J. Kim, and J.S. Choi, "Video retrieval based on Scene change detection in compressed streams," *IEEE Trans. Consumer Electronics*, Vol. 45, No. 3, pp. 932-936, 1999.  
 [8] B. Furht, S.W. Smoliar and H.J. Zhang, *Video and image processing in multimedia systems*, Kluwer Academic Publishers, 1995.

## 저 자 소 개



姜應寬(正會員)

1993년 2월 중앙대학교 전자공학과 졸업(공학사). 1995년 2월 중앙대학교 대학원 전자공학과 졸업(공학석사). 1995년 1월~1997년 5월 현대전자 정보통신연구소. 1997년 9월~현재 중앙대학교 첨단영상대학원 영상공학과 박사과정 재학 중. 주 관심분야는 동영상 부호화, 멀티미디어 데이터베이스, 영상통신, 컴퓨터비전 등임



宋浩根(正會員)

1991년 2월 중앙대학교 전자공학과 졸업(공학사). 1993년 2월 중앙대학교 대학원 전자공학과 졸업(공학석사). 1997년 2월 중앙대학교 대학원 전자공학과 졸업(공학박사). 1996년 3월~현재 한서대학교 컴퓨터정보학과 조교수. 주 관심분야는 패턴인식, 멀티미디어신호처리, 내용기반 영상검색 에이전트 시스템, 전자상거래를 위한 웹 호스팅 등임