

# Eigen Value 기반의 영상검색 기법

김진용\* · 소운영\* · 정동석\*

## Eigen Value Based Image Retrieval Technique

Jin-Yong Kim\* · Woon-Young So\* · Dong-Seok Jeong\*

### Abstract

Digital image and video libraries require new algorithms for the automated extraction and indexing of salient image features. Eigen values of an image provide one important cue for the discrimination of image content. In this paper we propose a new approach for automated content extraction that allows *efficient database searching using eigen values*. The algorithm automatically extracts eigen values from the image matrix represented by the covariance matrix for the image. We demonstrate that the eigen values representing shape information and the skewness of its distribution representing complexity provide good performance in image query response time while providing effective discriminability. We present the eigen value extraction and indexing techniques. We test the proposed algorithm of searching by eigen value and its skewness on a database of 100 images.

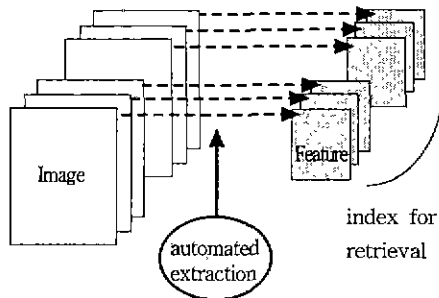
---

\* 본 연구는 한국 전자통신연구원 위탁연구의 일부 지원을 받아 수행한 것입니다.  
\* 인하대학교 전자공학과

# 1. 서론

## 1.1 영상의 특징을 이용한 indexing 방법

디지털 영상과 비디오의 수와 양에 있어서의 현격한 성장세는 이런 아무런 외부적 제한을 갖지 않는 영상물들의 색인화를 위한 도구 개발의 필요성을 느끼게 하였다. 이런 도구들은 기본적으로 영상 검색능력을 가지게 되었고, 이 때 영상들의 시각적 특징을 색인화하는 “내용기반 영상 질의(Content-Based Visual Query : CBVQ)” 기술이 사용되게 되었다. CBVQ의 목적은 영상인식(Image Understanding)의 도움을 받지 않고 시각정보에 대한 유사도 질의를 수행하는데 있어 유용한 구별자(discriminants)를 얻고 이용하는데 있다.



(그림 1-1) 영상 indexing의 개요

최근의 CBVQ에 관한 많은 연구들은 색, texture, 모양, 동작 그리고 위치적 정보 등과 같이 몇가지 특정한 시각적 차원에 그 초점이 맞추어져 왔다 [1~5, 8~10]. 이 들은 각각 <표 1-1>과 같이 영상을 통해 우리가 얻을 수 있는 정보를 이용한 것이다. 물론 이 특징들을 전부 다룰 수 있는 시스템이라면 더 할 나위 없이 이상적이겠지만 4>, 5>, 6>은 영상이해(Image Understanding)에 해당되는 것으로 현실적으로 1>, 2>, 3>(Image Recognition) 만으로도 어느 정도 만족할 만한 결과를 내고 있다. 하지만 수행결과와 품질을 높이기 위해 여러 정보를 동시에 이용하는 방법은 각각의 정보를 사용하기 위해 <표 1-2>와 같이 그에 따른 개별적

작업이 이루어져야 하는 번거로움이 존재하게 되었다.

<표 1-1> 영상을 통해 얻을 수 있는 정보

사람이 영상을 통해 얻는 느낌	객관화된 정보 종류
1> 밝다, 어둡다, 화사하다 등	color.
2> 날카롭다, 둥글둥글하다, 어떤 무늬다 등	shape, texture.
3> 단순하다, 복잡하다 등	complexity
4> 어떤 상징을 나타낸다	text, symbol.
5> 야하다, 중세풍이다, 세련되다, 멋있다 등	abstract domain concept.
6> 그 외 개개인의 환경에 따른 느낌 등	객관성이 결여된 concept.

<표 1-2> 영상정보의 종류에 따른 필요작업

이용 정보종류	필요한 작업
1> color.	color histogram 이용
2> texture.	co-occurrence matrix, s-f technique(Gabor filter, QMF filterbank 등) 이용
3> complexity	s-f technique 이용
4> shape	각종 shape descriptor(Moment, Chain code, Fourier 등) 이용

본 논문에서는 이런 문제점을 해결하기 위해 영상의 구별자(discriminants)로 영상의 eigen value 를 새롭게 도입하여 한번의 작업(계산)으로 영상의 모양과 복잡도를 얻어 CBVQ를 수행 할 수 있는 새로운 영상검색 시스템을 제안하였고 실제 인터넷상에서 서비스가 가능한 시스템을 완성 서비스를 실시 중이다. 부가적인 예기지만 이 시스템은 기존의 영상검색 시스템과는 달리 JAVA servlet을 사용하여 기존 CGI의 단점을 극복한 것도 큰 특징이라 하겠다.

## 2. Eigen based image analysis

비디오 라이브러리들을 다루는 도구들은 엄청난 양의 데이터를 계산하는 과정을 필연적으로 수반하게 된다. 따라서 이들을 실제 전체의 원 영상 정보 대신에 간략화된 데이터의 형태로 처리하는

것이 요구되어 왔다. 물론 이때의 간략화된 데이터는 원 영상의 정보를 충분히 포함하는 정도이어야 하고, 실제로 이런 간략화된 데이터로부터 효과적인 구분 특징을 뽑아 낼 수 있어야만 한다. 영상내의 오브젝트의 외형상의 유사도를 효과적으로 측정하기 위해서는 그 영상내의 오브젝트를 표현하는데 있어 먼저 어떤 특징이 지배적인가를 결정하는 것이다. 이 때 특징을 추출하는 데 있어 사용되는 표준 선형적 방법으로 주성분 분석(principal components analysis)법이 있다[15, 3]. 이 주성분 분석법은 데이터의 차원을 줄이기 위해 상관된 랜덤 변수(correlated random variables)를 분석함으로써 이루어지는 수학적 기법이다. 즉 간략화는 단순히 데이터의 첫 주성분 몇 개를 뽑아냄으로써 이루어지게 되고, 이 때 주성분은 오브젝트를 구분하는 데 있어 가장 관련되는 특징들을 지니게 되는 것이다.

이를 적용하기 위해 먼저 영상을 공간 차원(spatial domain)에서 인자  $q$ 만큼의 부표본화(subsampling)하여 데이터를 줄이게 된다. 부표본화는 다음과 같이 처리되고, 부표본화된 영상의 표현은 식 (2-1)과 같다.

$$I_s(x, y) = I(qx, qy) \quad (2-1)$$

, where  $1 \leq x \leq X$  and  $1 \leq y \leq Y$

$f$ 를  $N \times 1$  부표본화된 영상  $I$ 의 열벡터(column vector)로 나타내고,  $\mu$ 를 평균벡터(mean vector)라 하면, empirical covariance matrix는 다음과 같이 얻어 질 수 있다.

$$S = \frac{1}{M} \sum_{y=1}^Y \phi_y \phi_y^T$$

where  $S$  : empirical covariance matrix,  $\phi_y = f_y - \mu$   
 $f$  :  $N \times 1$  column vector  
 $\mu$  : mean vector

이 때 우리는  $S$ 에 대하여 unique한  $Y$ 개의 Orthogonal eigen vector :  $[q_1, q_2, q_3, \dots, q_Y]$  와 그들과 관련된 eigen value  $[\lambda_1 > \lambda_2 > \lambda_3 > \dots > \lambda_Y]$  를 계

산할 수 있다. 또 이 때 첫  $K$ 번째까지의 선형 결합(linear combination)  $Q = [q_1, q_2, \dots, q_K]$ 는 부표본화된 영상의 대부분의 영상 정보를 함축하며 거친 해상도(coarse resolution)로 원영상의 공간(space)으로 확장(span)한다.

따라서 영상은 크기순으로  $K$ 개의 eigen value에 상응하는 eigen vector로 투사되고 다음 식과 같이 eigen-image 표현으로 변형된다.

$$p = Q^T \phi = [p_1, p_2, \dots, p_K]^T \quad (2-3)$$

이 때  $p_1, \dots, p_K$ 를 주성분(principal components)라 하고, 벡터  $p$ 는 영상의 간략화된 정보를 지니게 된다.

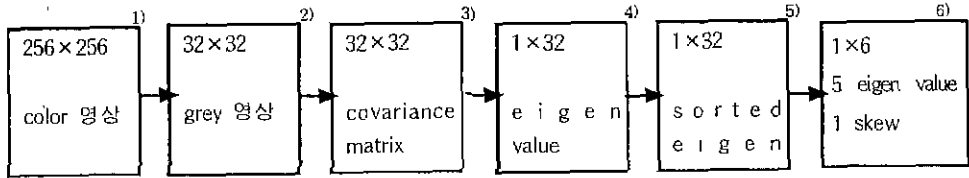
### 3. 영상의 특징으로서의 eigen value

#### [정의 3.1]

symmetric matrix  $A$ 에 대하여  $Ax = \lambda x$ 을 만족시키는  $0$ 이 아닌 vector  $x$ 가 존재할 때,  $\lambda$ 를 matrix  $A$ 의 eigen value,  $x$ 를 eigen vector라 한다.

이 때 matrix  $A$ 를 영상이라고 보면 이 영상에는 자신만이 지니는 eigen value와 eigen vector를 갖는다는 것을 알 수 있다. 즉 이 영상의 고유한 특징(distinct feature)으로서 사용될 수 있는 것이다. 물론 일반 영상은 symmetric이 아니기 때문에 그 영상의 정보를 대부분 포함하면서 symmetric인 조건을 만족시키기 위해 식 (3-3)과 같이 영상의 Empirical Covariance Matrix를 이용한다. 이는 영상 matrix의 행간의 통계적 관계를 이용하는 것으로 전반적인 영상정보를 포함하게 되는 반면에 정보를 얻는 과정에 있어 방향성을 타게 되어 후에 eigen value가 영상의 translation에는 별 영향을 받지 않으나 rotation에는 영향을 받는 원인으로도 작용한다.  $M \times M$  영상  $f$ 에 대하여

$$f_{ij} = \begin{bmatrix} f_{00} & f_{01} & \dots & f_{0M} \\ f_{10} & f_{11} & \dots & f_{1M} \\ \dots & \dots & \dots & \dots \\ f_{M0} & f_{M1} & \dots & f_{MM} \end{bmatrix} \quad (3-1)$$



(그림 3-1) eigen value 추출과정

$$\sigma_{ij} = E[(f_i - \mu_i)(f_j - \mu_j)], \quad (0 \leq i, j \leq M)$$

the variance between  $f_i$  and  $f_j$ , (3-2)

$$C_f = \begin{bmatrix} \sigma_{00} & \sigma_{01} & \dots & \sigma_{0M} \\ \sigma_{10} & \sigma_{11} & \dots & \sigma_{1M} \\ \dots & \dots & \dots & \dots \\ \sigma_{M0} & \sigma_{M1} & \dots & \sigma_{MM} \end{bmatrix} \quad (3-3)$$

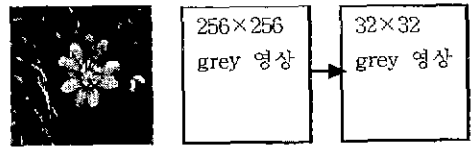
$f_x$ : the  $x$ -th row of the image  $f$   
 $\mu_x$ : mean vector of the  $x$ -th row of the image  $f$

### 3.1 추출과정

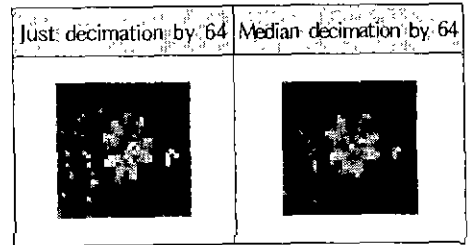
(그림 3-1)은 실제 영상으로부터 eigen value를 추출하는 과정으로

- 1) grey영상화후 8x8 block에 대한 median down sampling and histogram equalization
- 2) empirical covariance matrix 계산
- 3) Jacobian transform을 이용한 eigen value 추출
- 4) 32개의 eigen value들을 내림차순으로 정렬
- 5) 32개로 이루어진 eigen value distribution에 대한 왜도(skew)를 구하고, 4분위수 위치의 eigen value 5개를 이 영상의 특징으로 색인한다.

Median down sampling을 하는 이유는 앞서 언급한 정보의 간략화를 위해 영상정보의 redundancy를 줄이고, 계산의 효율성을 높일 수 있고 또 median filter를 통해 (그림 3-2)와 같이 object의 coarse segmentation 효과를 얻을 수 있기 때문이고 histogram equalization이 필요한 이유는 principal components analysis 자체가 주성분에 크게 의존하다 보니 밝기(Luminance or Intensity)의 영향에 크게 좌우되어 똑 같은 영상이라도 feature가 달리 나오는 약점을 보완하기 위해 밝기



원본 영상



(그림 3-2) 제안된 median down sampling을 통한 object의 coarse flattening

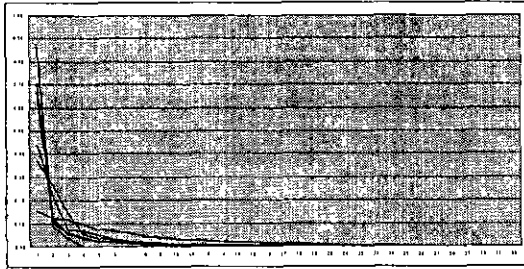
보정 차원으로 적용되는 과정이다.

Empirical covariance matrix로 부터 Jacobian Transform을 통해 eigen value를 얻어낸다[14]. 이후 얻어진 eigen value를 내림차순으로 정렬하여 4분위수에 해당하는 eigen value만을 영상의 특징으로 취한다. Hotelling transform[6]을 통해서도 알 수 있듯이 5개의 eigen value가 전체 영상의 대략 90% 이상의 정보를 나타내기 때문에 1차적으로 얻은 32개의 eigen value중에서 크기순으로 첫 5개만으로 충분히 영상의 특징을 나타낼 수 있지만 전체적인 분포의 모양을 고려할 때 4분위수 위치의 값을 영상의 특징으로 추출하게 된다.

### 3.2 정량화 및 그 이용

일반적으로 위 추출과정을 통해 얻어진 32개의 eigen value들을 내림차순으로 정렬했을 때, 그 분

포가 (그림 3-3)과 같이 대략 Rayleigh distribution 이 됨을 알 수 있다.



(그림 3-3) 영상의 eigen value 분포의 예

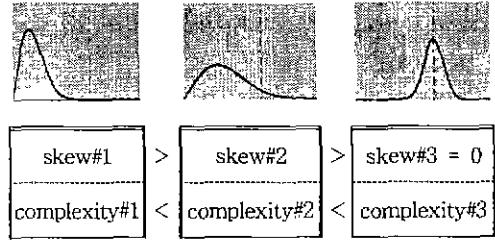
3.2.1 모양의 표현

앞서 언급했던 principal components analysis를 통해 영상내 object를 표현 할 수 있음을 알았다. 하지만 그 방법은 eigen value와 eigen vector를 계산해야 하는 상당한 계산량을 필요로 함으로써 실제 구현에 많은 문제가 있다[14]. 본 논문에서는 eigen value만으로 특징을 제한함으로써 영상의 방향성에 대한 제약을 받지만 상대적으로 계산량을 크게 줄일 수가 있었다. 또 eigen vector와 무관하게 eigen value자체의 크기 만으로도 영상내 object 모양을 다음과 같이 나타낼 수 있다. eigen value가 크다는 의미는 그 eigen value와 연관된 방향(eigen vector)성분이 크다는 뜻이고 이를 통해 image내 object의 개략적인 shape을 나타낼 수 있기 때문이다.

3.2.2 skew를 통한 복잡도 구현

또 이를 이용하여, eigen value가 전 방향성분 (본 논문의 경우는 32개)에 걸쳐 고르게 분포한다면 object의 shape가 복잡하게 되고 한 방향으로 치우친다면 단순하게 되어 복잡도를 나타낼 수 있는 것이다. 이를 정량화하기 위해 식 (3-4)와 같이 distribution의 형태를 통한 분석으로 왜도(skew)를 도입하여 복잡도를 계산하게 된다.

$$skew = \frac{n}{(n-1)(n-2)} \sum \left( \frac{x_i - \bar{x}}{s} \right)^3 \quad (3-4)$$



(그림 3-4) skew와 복잡도와의 관계

(그림 3-4)를 통해서 알 수 있듯이 분포의 모양이 정규 분포인 경우 eigen value가 골고루 퍼져 있게 되어 skew는 0이되고 이는 영상의 principal components가 몇 개의 성분이 아닌 전 성분에 고루 퍼져 있다는 뜻으로 영상이 복잡하다는 의미이고, 마찬가지로 왜도가 심한 경우 principal components가 몇 개의 성분에만 집중하는 단순한 형태의 영상이 된다는 의미다. 실제 영상에 대해 적용한 결과를 (그림 3-5)에 나타내었고 영상의 eigen value 분포의 skew는 다음과 같이 복잡도를 잘 표현한다.



(그림 3-5) 영상의 eigen value분포의 skew와 영상의 복잡도와 관계 실예

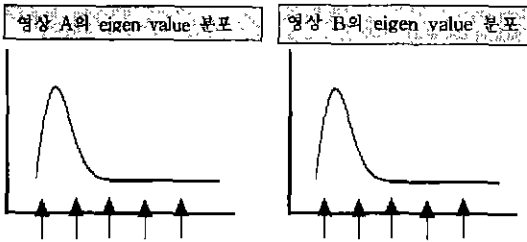
3.2.3 유사거리(similarity distance) 계산

exponentially weighted quartile distance =

$$\begin{aligned} & [ (Quartile(A, 0) - Quartile(B, 0)) * e^{-8} \\ & + (Quartile(A, 1) - Quartile(B, 1)) * e^{-6} \\ & + (Quartile(A, 2) - Quartile(B, 2)) * e^{-4} \\ & + (Quartile(A, 3) - Quartile(B, 3)) * e^{-2} \\ & + (Quartile(A, 4) - Quartile(B, 4)) * e^0 ] * 10^6 \end{aligned} \quad (3-5)$$

where  $A, B$  denotes two eigen value distributions respectively, and  $Quartile(A, n)$  denotes  $n$ -th quartile of  $A$ .

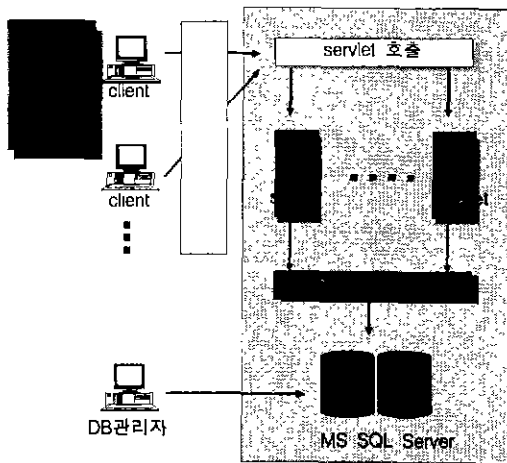
두 영상의 유사도를 측정하기 위해 제안한 *exponentially weighted quartile distance*는 기본적으로 두 distribution의 Euclidean distance를 구하는데, 그 크기에 있어서 아주 작지만 비교시 실질적으로 변별성을 나타내는데 중요한 역할을 하는 detail information을 강조하기 위해 exponential weighting을 주었고, 전 eigen value 중 4분위수에 해당하는 5개의 eigen value들간의 distance를 두 영상의 유사거리로 취했다.



(그림 3-6) 두 영상의 유사거리 계산시 비교 위치

## 4. 실험결과 및 고찰

### 4.1 실험환경



(그림 4-1) 제안된 알고리즘을 적용한 영상 검색 시스템.

제안된 알고리즘을 적용하여 제작한 시스템은 크게 세 부분으로 이루어져 있다

#### 4.1.1 Content based feature extractor

DB관리자에 해당하는 부분으로 특징 추출과정 대로 영상에서 특징을 추출하여 데이터베이스를 구축한다. 실험에서는 C++ code를 이용하여 MS SQL server용 DB를 생성한다. 이 때 생성되는 DB는 기본적으로 1개의 file, 1개의 table에 각 record는 9개의 field로 구성되게 된다.

〈표 4-1〉 데이터베이스의 구성

ID	picturename	eig #1	eig #8	eig #16	eig #24	eig #32	skew	distance
1	flowr001	###	###	###	###	###	###	
2	flowr002	###	###	###	###	###	###	
:	:	:	:	:	:	:	:	

#### 4.1.2 Query generator

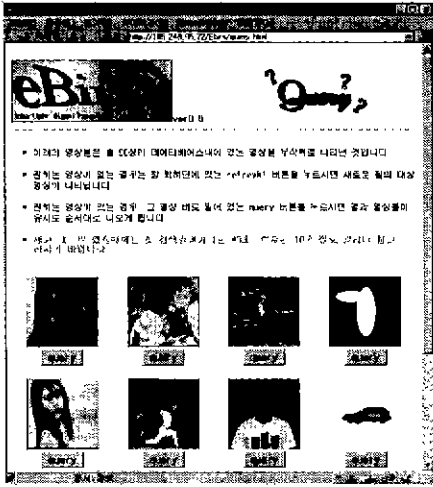
Web browser 사용자로부터 질의 영상을 받고 그 영상의 특징을 뽑아 처리용 query data를 발생 시켜 주는 부분으로 JAVA applet과 HTML code 로 작성한다.

#### 4.1.3 Query manipulator

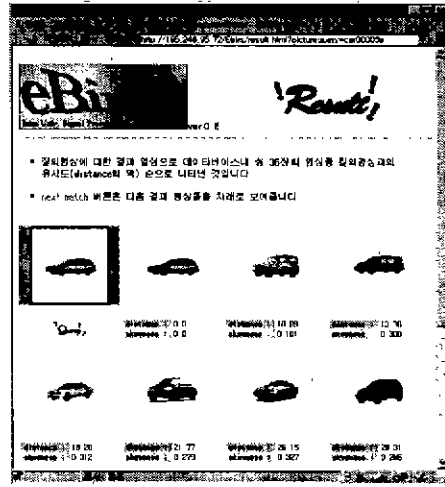
Query generator로부터 발생한 query data를 받아 기존의 구축된 DB내의 영상특징들과의 유사도를 계산하여 가까운 순서로 출력시키는 부분으로 JAVA servlet으로 작성한다.

## 4.2 실험결과

(그림 4-2, 4-3)은 실제 본 논문에서 제안한 영상검색 시스템에 Web browser로 접속하여 자동차영상을 질의영상으로 주고 결과를 받는 것을 보여주고 있고 (그림 4-4)는 참고적으로, 결과로서 나오게된 자동차 영상들의 eigen value분포를 나타낸 것이다.



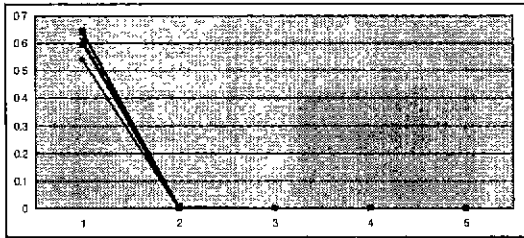
(그림 4-2) 질의 입력 화면



(그림 4-3) 질의에 대한 결과 화면

Web browser로부터 자동차영상을 질의 영상으로 입력 받음

결과 영상을 Server로부터 받음



(그림 4-4) 자동차 영상들의 eigen value분포

(그림 4-5)의 경우 skew를 고려하지 않은 상태에서 단순히 eigen value의 유사거리만을 나타낸 결과로 사용자로부터 조금은 납득하기 어려운 영상이 존재하지만 일관성이 있음을 알 수 있다. 특히 eigen value가 실제적으로 영상에 어떤 식으로 적용이 되는 지는 (그림 4-5(b))를 통해 알 수 있듯이 영상이 object개념으로 segment가 되면 그의 상징적인 모양(shape)으로도 표현할 수 있게 되는 것이다. 또 이때 skew를 적용하면 (그림 4-6)과 같이 그 납득하기 힘든 영상들이

뒤로 처지게 되어 주어진 질의에 대한 결과의 정확도를 향상시킬 수 있고 후보영상의 범위를 1차적으로 줄여 줌으로써 질의 반응시간을 단축시키는 부수적인 효과도 얻을 수 있다.

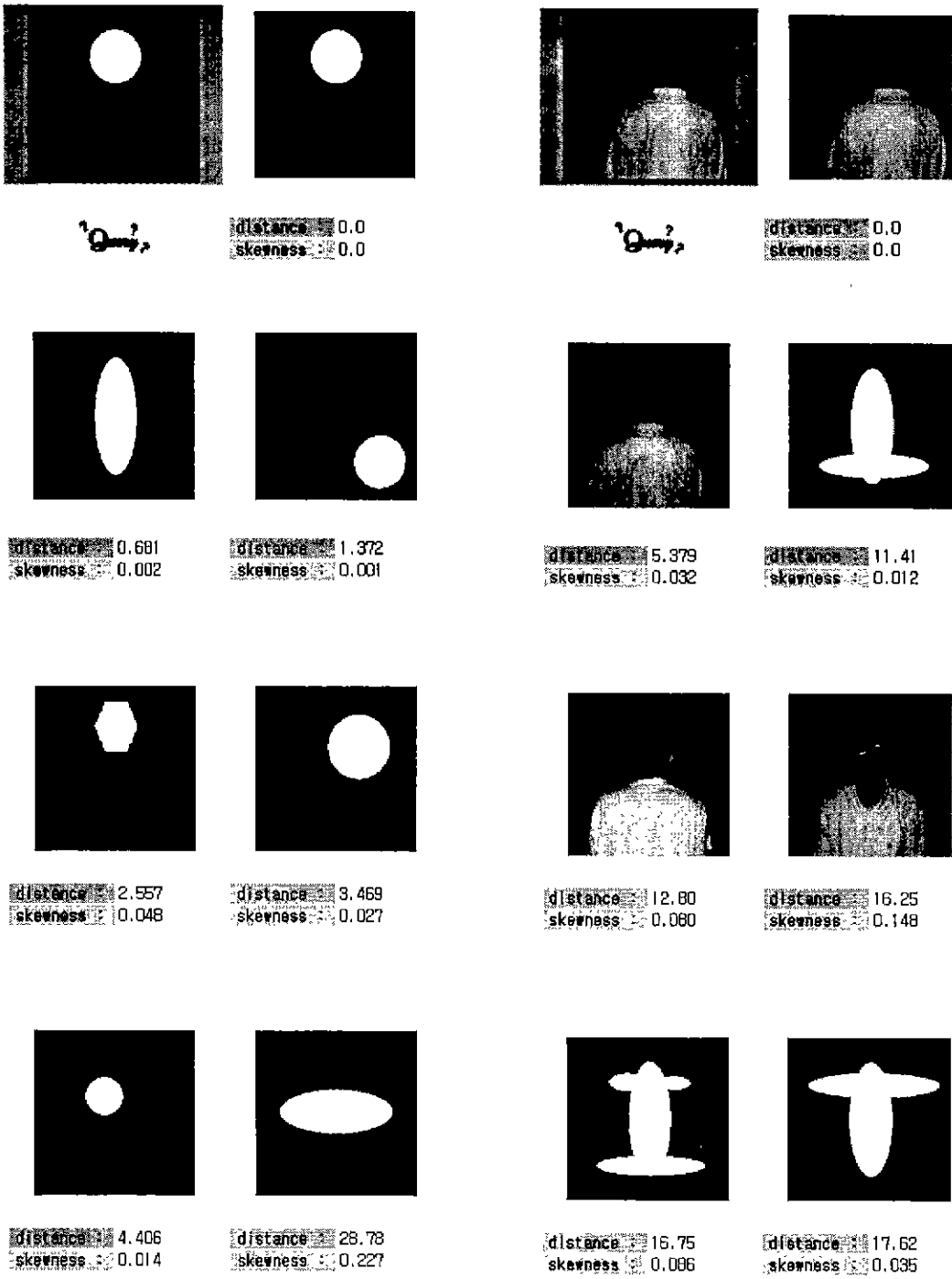
### 5. 결 론

1) 단지 6개의 feature로 전체적인 image정보를 표현 할 수 있어 data를 효과적으로 이용할 수 있고, 한번의 eigen value 계산으로 image의 복잡도와 모양을 나타낼 수 있다. 이 때 결과의 정확도와 반응시간간의 보상을 통해 feature의 개수를 조절할 필요가 있다.

2) Eigen value가 밝기의 변화에 민감하기 때문에 histogram equalization과 같은 보정이 필요하고 또 영상의 translation에는 변화가 없지만, rotation과 scaling에는 변하기 때문에 이에 대한 고려가 필요하다. 또 풍경화와 같은 object개념이 적용되기 힘든 영상의 경우는 제한할 필요가 있다.

3) 인물의 동작영상을 실험으로 한 결과를 통해 볼 때 일반 영상보다는 연속된 동영상의 경우에 적용할 경우에 장면전환에 관한 정보를 얻을 수 있어 이 분야로의 확대도 가능하리라 생각된다.

4) 영상의 정보라는 것이 색, 모양, 질감이 대부분인 것을 감안할 때 향후 색과 질감을 추가할 경우 정확도에 있어서 괄목할 만한 결과가 나오리라 기대할 수 있다.

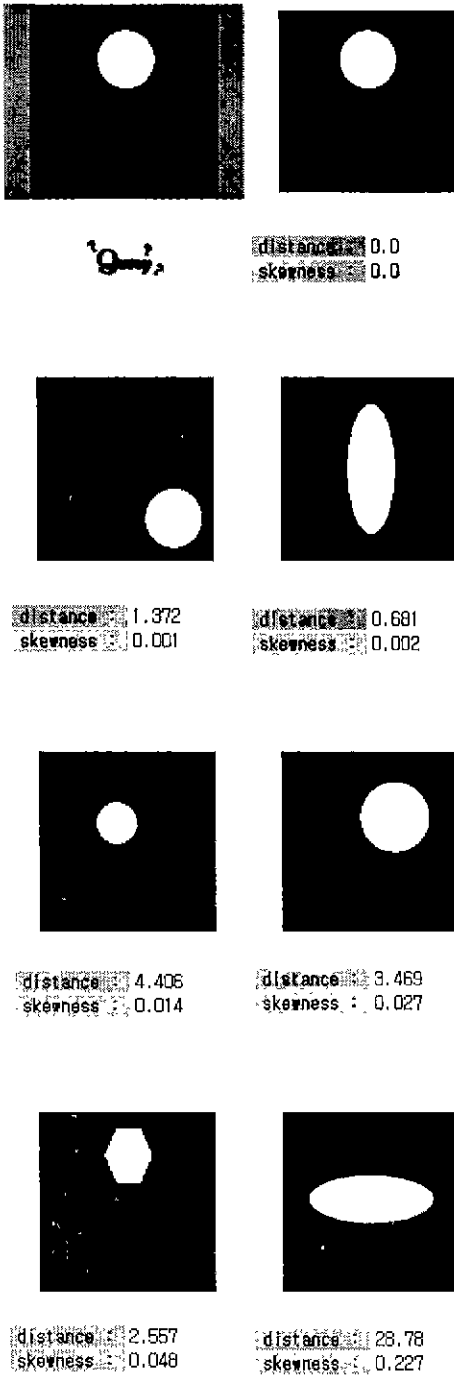


(a)

(b)

(그림 4-5) eigen value만을 이용하여 영상 검색을 수행한 결과





(그림 4-6) eigen value와 skew를 이용하여 영상 검색을 수행한 결과

## 참고 문헌

- [1] Keesook J. Jan, Ahmed H. Twefik, "Eigen-Image Based Video Segmentation and Indexing," *IEEE International Conference On Image Processing*, vol.II, pp.538-541, 1997.
- [2] E. Ardizzo, V. D. Gesu, "Content based indexing of Image and Databases by Global and Shape Features," *IEEE Proceeding of ICPR*, pp.140-144, 1996.
- [3] John R. Smith, Shin-Fu Chang, "AUTOMATED TEXTURE FEATURE SETS FOR IMAGE RETRIEVAL," *IEEE ICASSP 96* pp.2239-2242, 1996.
- [4] John R. Smith, "Integrated Spatial And Feature Image System," *the degree of Doctor of Philosophy in the Graduate School of Arts and Sciences*, Columbia University, 1997.
- [5] Roger Mohr, Patric Gros, Cordelia Schmid, "Efficient Matching with Invariant Local Descriptors," *Korea-France Workshop on Multimedia*, pp.1-18, 1998.
- [6] Rafael C. Gonzalez, Richard E. Woods, "THE HOTELLING TRANSFORM, REPRESENTATION AND DESCRIPTION" "Moments", *Digital Image Processing*, pp.148-156, pp.514-518, 1992.
- [7] 김진용, 소운영, 경동석, "Eigen value를 이용한 영상검색시스템," '98 신호처리 합동학술대회 논문집, pp.997-1000, 1998.
- [8] M. J. Swain, D. H. Ballard, "Color Indexing," *Int. Journal of Computer Vision*, pp.11-32, 1991.
- [9] W. Niblack, R. Barber, The QBIC Project : Querying Images by content using color, texture, and shape. In storage and Retrieval for Image and Video Database I, Vol., 1908 of *SPIE Proceeding series*, Feb 1993.

- [10] Kim yongsung, "Content-Based Trademark Retrieval System Using Visually salient Feature," *Proceedings of IEEE International Conference on CVPR 97*, 1997.
- [11] Patrick C Hew, "Recognition of Printed Digits using Zernike or Orthogonal Fourier-Mellon Moments," *the Department of Mathematics*, 1997.
- [12] Jimmie Gilbert, Linda Gilbert, "Eigenspaces and Similanty," *Linear Algebra And Matrix Theory*, pp.222-231, Academic Press, 1970.
- [13] Gilbert Strang, "Eigenvalues and Eigenvectors," *Linear Algebra And Its Application*, pp.243-319, Harcourt Brace Jovanovich, 1976.
- [14] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery, "Eigen Systems," *Numerical Recipes in C*, pp.456-493, Cabridge University Press, 1988.
- [15] Jain, "THE KL TRANSFORM," *FUNDAMENTALS OF DIGITAL IMAGE PROCESSING*, pp.163-175. Prentice Hall, 1989.
- [16] Thomas P. Weldom, William E. Higgins, Dennis F. Dunn, "Gabor filter design for multiple texture segmentation," *SPIE*, 1996.

#### 저자소개



#### 김진용

저자 김진용은 인하대학교 전자공학과를 졸업하고, 공학석사를 취득하였으며, 일본마이크로소프트 Commerce Server Team에 소속되어 있습니다. 주 관심분야는 IMT2000, 전자상거래 입니다.



#### 소운영

공동저자 소운영은 인하대학교 전자공학과를 졸업하고, 현재 Multi-Media Lab에서 석사과정 중입니다. 주 관심분야는 의료 영상 처리 분야입니다



#### 정동석

공동저자 정동석은 서울대학교 전기공학과를 졸업하고, 미국 Virginia 주립 공과대학 전자공학과 공학석사, 박사학위를 취득하였으며 현재 인하대학교 공과대학 전자공학과 교수로 재직하고 있으며 주 관심분야는 영상처리, 컴퓨터 비전, 멀티미디어 정보처리입니다.