

MPEG-7 시각 기술자와 해마 신경망을 이용한 내용기반 검색

Content-Based Retrieval using MPEG-7 Visual Descriptor and Hippocampal Neural Network

김 영 호, 강 대 성*
(Young Ho Kim and Dae-Seong Kang)

Abstract : As development of digital technology, many kinds of multimedia data are used variously and requirements for effective use by user are increasing. In order to transfer information fast and precisely what user wants, effective retrieval method is required. As existing multimedia data are impossible to apply the MPEG-1, MPEG-2 and MPEG-4 technologies which are aimed at compression, store and transmission. So MPEG-7 is introduced as a new technology for effective management and retrieval of multimedia data. In this paper, we extract content-based features using color descriptor among the MPEG-7 standardization visual descriptor, and reduce feature data applying PCA(Principal Components Analysis) technique. We model the cerebral cortex and hippocampal neural network in engineering domain, and learn content-based feature vectors fast and apply the hippocampal neural network algorithm to compose of optimized feature. And then we present fast and precise retrieval effect when indexing and retrieving.

Keywords : MPEG-7 standardization visual descriptor, hippocampal neural network algorithm

I. 서론

디지털 기술의 발달로 수많은 멀티미디어 형태의 데이터들이 다양하게 사용되고 있으며, 사용자가 이를 효율적으로 이용하기 위한 수요 또한 늘어나고 있다. 사용자가 원하는 정보를 빠르고 정확하게 전달하기 위해선 효과적인 검색 방법이 필요하다. 기존의 멀티미디어 데이터는 압축, 저장, 전송에 초점이 맞춰진 MPEG-1, MPEG-2, MPEG-4 기술로는 적용이 불가능하기 때문에 멀티미디어 데이터의 효율적인 관리 및 검색을 위한 새로운 기술로 MPEG-7이 등장하게 되었다. MPEG-7은 멀티미디어 정보를 검색하기 위한 다양한 형태의 표준화된 특징 표현 기술의 집합으로 "multimedia content description interface"라 할 수 있다[1,2].

기술자(descriptor)는 멀티미디어 정보 내용상의 구별되는 특징(feature)을 표현하기 위한 기본이 되는 기술이다. 여기서의 특징이란 색상(color), 질감(texture), 형태(shape), 움직임(motion) 등이 될 수 있다. 이러한 기술자를 이용해서 사용자가 원하는 멀티미디어 정보를 빠르고 효과적으로 검색할 수 있다. 본 논문에서는 MPEG-7 시각 기술자중 HMMD 칼라 모델과 edge 히스토그램 기술자(EHD)를 이용하여 내용기반 특징을 추출하고 PCA 기법을 적용하여 특징 데이터를 축소 시켰다[3,4].

특징 데이터를 색인 및 검색을 위해서 해마 신경망에 적용하였다[5]. 공학부분에서 많이 응용되고 있는 신경망은 뇌의 학습-기억 활동을 모태로 다양한 분야에 응용되고 있다. 신경망에서의 학습은 실제 뇌 안에서 뉴런과 뉴런이 각각 시냅스로 연결되어 정보 전달 및 처리 하는 과정으로

볼 수 있다. 따라서 신경망을 단편적으로 고려했을 때 신경망의 학습과정은 실제 뇌에서의 시냅스를 표현하는 가중치 조절과 긴밀히 연관되어 있음을 알 수 있다. 시냅스의 기능과 구조를 신경망에 반영하는 것은 실제 뇌 안에서 발생하는 정보처리 과정을 설명하기에 적절하다. 시냅스를 통한 신경전달물질간의 이동은 시냅스에서 일어나는 신경물질의 전도과정(synaptic transmission)으로서 뉴런과 뉴런간의 의사소통 수단으로 이용된다. 시냅스가 갖는 신경 세포간의 의사소통 기능은 전기적 신호를 화학적 신호로 변화시키는 중간 단계이며, 실제 신경 세포 내에서 일어나는 생화학적 기작을 설명하는 데에 적합하다. 자극이 주어졌을 때 전 시냅스에서 정보를 내포하고 있는 분자단위의 물질들이 활성화 되어 후 시냅스로 진행되는 동안 분자 혹은 기질간의 다이내믹스는 연속적이며, 병렬적으로 발생한다. 헤브의 학습방법을 바탕으로 생성된 기존의 신경망들은 실제 뇌 모델에서의 그것과 차이가 있다. 뇌에서는 실제로 해당 패턴 하나에 필요한 흥분 세포들의 정보만 저장된다. 일반적인 신경망 구조를 갖는 체제에서는 필요한 특징들만 저장할 수 없다.

본 논문에서는 특징벡터를 효율적으로 이용하기 위해 단순히 백터거리 비교 방법을 쓰지 않고, 제한한 해마 신경망 알고리즘을 적용 시킨 후 인식을 하게 된다. 영상에서 추출한 특징 배열들을 해마 신경 구조의 3단계 과정을 거쳐 단기 기억과 장기 기억으로 나누고 오인식이 일어나는 단계에서 동적으로 뉴런의 개수를 증가시켜서 새로운 기억으로 만들며, 재학습의 단점을 극복하여 검색 시간을 줄이고 높은 검색률을 보였다.

본 논문의 구성은 다음과 같다. 2장에서는 특징 추출 방법을 소개하고, 3장에서는 해마 신경망을 설명하고, 공학적으로 모델링 한 과정을 소개한다. 4장에서는 실험 결과를

* 책임저자(Corresponding Author)

논문접수 : 2005. 5. 2., 채택확정 : 2005. 11. 7.

김영호 : 동아대학교 전자공학과(newkyh@hotmail.com)

강대성 : 동아대학교 전자컴퓨터공학부(dskang@daunet.donga.ac.kr)

분석하였고, 5장에서는 결론을 맺고 추후 연구 방향을 제시한다.

II. 특징 추출

1. HMMD 컬러 모델

공간 HMMD 컬러 모델은 무색 범위(achromatic region)와 컬러 범위(chromatic region)에서 5 공간으로 나누어진다. 여기서 무색 범위는 밝기 요소 (brightness factor)를 기반으로 한 양자화이고, 컬러 범위는 4개 예술적인 구성들(artistic components) 즉, "hue", "tint", "tone" 그리고 "shade"를 기반으로 양자화 되었다. HMMD 컬러 모델에는 5개의 파라미터를 가지고 있다. Hue는 Hue 범위 0°에서 360°에 의해 표현된다. 각이 증가하면, H는 빨간색(0°=360°), 노란색(60°), 초록색(120°), 파란색(240°)으로 변화된다. Max는 검은색의 양을 말하며, 색의 농도(shade)를 준다. Min은 흰색의 양을 말하며, 색조(tint)를 준다. Diff는 단일(pure) 색에 가깝고, 그레이의 양을 말하며, 명암(tone)을 준다. Sum은 색의 밝기를 계산한 것이다. 여기서 Hue, Max 그리고 Min 또는 Hue, Diff 그리고 Sum인 3개 파라미터만으로도 컬러 공간의 분포를 해석하기에 충분하다. 그림 1은 HMMD 컬러 모델의 좌표계이다. 여기서 Hue는 Diff 축에 직각으로 있는 수직 축 둘레의 각이고, Diff와 Sum은 Max와 Min의 차와 합에 의해 결정되는 보조 파라미터이다. HMMD 컬러 모델은 RGB 컬러 모델을 통하여 얻을 수 있다. 여기서 Max와 Min 그리고 Diff, Sum은 0에서 1사이의 값이고, H는 0에서 360 사이에 존재한다.

2. Edge 히스토그램 기술자(EHD)

Edge 히스토그램 기술자(EHD)는 전체 영상 공간을 4X4의 16개 부분 영상으로 분할한 후, 각 영역별 edge 히스토그램의 분포를 나타낸다. 여기서 부분 영역(semi-global) 히스토그램은 16개의 부분 영상에 대하여 화면의 가로, 세로 방향에 대한 8종류(가로 방향 4, 세로 방향 4)의 부분 영역을 정의한다. 영상 공간을 겹치지 않도록 작은 정방형 블록으로 나누며, 특히, 영상의 크기에 비례한 블록 크기를 결정함으로써 전체 영상에 대하여 고정된 개수의 블록을 발생시키도록 하였다. 이렇게 함으로써 영상 내 edge를 추출할 블록의 개수가 입력된 영상의 크기와 무관하게 일정한 값으로 유지되고, 한편, 블록 내 edge 검출을 쉽게 하기 위하여 블록의 크기(block_size)는 2의 배수가 되도록 조정하였다. 그리고 이것을 계산하기 위해서 16개의 부분 영역들은 다시 이미지 블록들로 나뉜다. 여기서 원하는 정확도에 따라 이미지 블록의 개수가 정해지지만 6X6(36)개의 블록으로 나누는 것이 보통이다. 각각의 블록을 2X2로 나누어 평균을 낸 후 edge의 방향성을 구하기 위해서 마스크를 씌운다.

5개의 bin 중에서 최대값이 문턱치를 넘으면 그 bin의 count를 하나씩 증가시킨다. 주어진 영상의 edge 히스토그램 기술자(EHD) 값은 전체 영상에 대한 각 bin의 평균값을 사용한다. 각 부분 영역에 대하여 edge 종류별 누적 값을 사용하여 특징정보를 생성한다. 여기서 edge들은 5개의 그룹으로 나뉘는데, 수직, 수평, 45도, 135도, 등방성(일정한

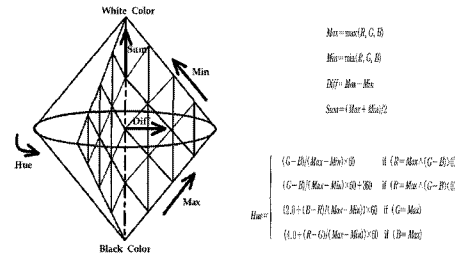


그림 1. HMMD 컬러 모델의 좌표계.

Fig. 1. Double cone representation of the HMMD color space.

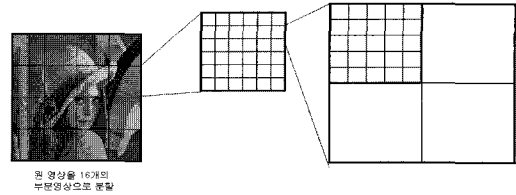


그림 2. Edge 히스토그램 기술자(EHD)의 처리 과정.

Fig. 2. Definition of sub-image and image-blocks.

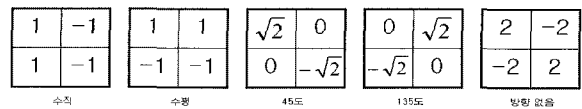


그림 3. Edge 히스토그램 기술자(EHD)를 위한 마스크들.

Fig. 3. Mask for edge histogram descriptor.

방향이 없음)이다. 각 부분 영역마다 특징정보는 5가지 edge를 나타내는 bin(bin)으로 이루어졌으므로 각 부분 영상마다(4X4=16개) 카운트된 특징정보를 해당 위치의 5가지 히스토그램 bin의 업데이트에 사용한다.

따라서 부분 영역 히스토그램이 가지는 총 bin의 수는 16X5=80 개이다. 따라서 전체 edge 히스토그램 기술자(EHD)은 전체 영상 내 발생된 edge의 전부를 5가지 종류에 따른 히스토그램을 구성한 것이다.

III. 특징점 학습 및 검색

해마의 기능 중 가장 중요하고 강력한 것은 단기 기억에서 장기 기억으로의 확장에 있다. 해마의 역할은 방금 저장된 정보를 기억하고 필요한 정보와 그렇지 못한 정보를 분류하는 것이다. 필요한 정보의 경우, 해당 뉴런에 저장해두고 동일한 패턴이 입력되면 인식하게 되는 것이다. 단기 기억과 장기 기억으로 나누는 방법에 있어서 호감도 측정은 통계적 패턴 인식에 사용되는 방법을 사용하면 통계적으로 많이 나타나는 값들에 대해서 호감도를 높게 부여하는 것이다. 이러한 특징을 해마 신경망에도 도입한다면, 입력되는 중간에 입력받는 값들의 가중치를 사용하여 단기 기억을 할 것인지 장기 기억을 할 것인지를 결정할 수 있다. 이러한 결정 값을 해마 기억 장소 분류 기준 값으로 사용할 수 있다. 전체 데이터 개수 대비 해당 뉴런의 누적 개수가 문턱치를 만족하면 장기 기억 장소로 저장시키도록 한다.

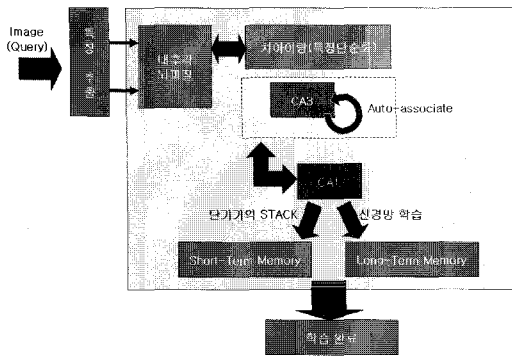


그림 4. 해마 신경망 학습 모델.

Fig. 4. Hippocampal neuron model.

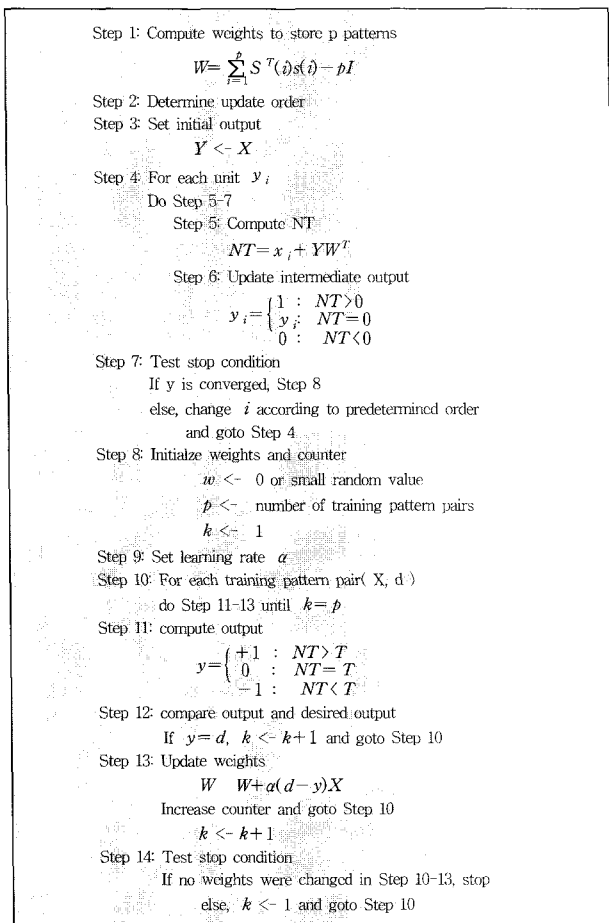


그림 5. 해마 신경망 알고리즘.

Fig. 5. Pseudo code of HNMA.

- 내추각뇌피질 : 해마와 신피질 사이에서 인터페이스를 구성하고 해마 신경망 모델의 입력과 출력을 나타낸다.
- 치아이랑 : 내추각뇌피질에 직접 연결되는 치아이랑 구조는 그 모델에 동일한 객체의 다양한 특징 성분들을 과거 입력 패턴에서 동일성을 판단하여 특징을 단순화시킨다. 패턴 평균값에서 편차율 범위의 임계수치를 넘으면 1, 그렇지 않으면 -1로 하여 특징의 동일성을 이진화시킨다.
- CA3 : 어떤 사건을 기억해낼 때에는 연상 작용을 반복함

으로써 보다 나은 결과를 얻을 수 있으며, 이러한 순환 개념을 도입한 것이 순환 연상 메모리이다. 순환 연상 메모리는 그림 4에서 보는 바와 같이 출력이 입력에 재환 되는 동질 연상 메모리이다. CA3 영역은 치아이랑에 연결되고 이 후자 구조로부터 단순화된 정보를 받는다. CA3 영역에서는 자기 연상을 하며, 노이즈가 있는 특징 정보를 더욱 분별력 있게 정돈하는 역할을 한다. 과거에 학습된 특징들과 차이가 큰 특징이 입력 될 경우에는 새로운 패턴으로 분류한다.

- CA1 : CA1 영역은 CA3 영역에 연결되고 정보 처리의 마지막 단계를 나타낸다. 자기 연상된 CA3의 정보를 단층 신경망에 학습 시키며, 장기 및 단기 기억을 결정한다. 학습된 가중치에 반응하여 출력 값이 맞지 않을 경우 단기 기억으로 패턴을 분류하여 저장하는 기능을 한다.

IV. 실험결과

1. 실험 방법

내용 기반 검색을 위한 실험 시스템 사양은 다음과 같다.

- HP xw8000 dual 2 GHz Xeon 워크스테이션
- Windows XP Professional with Service Pack 1
- 1 GB 시스템메모리
- 120 GB internal IDE drive

본 논문은 192x144 크기의 영상 1307개의 컬러 (HMMD), 질감(EDH)을 특징으로 한다. 특히 컬러모델은 차원을 줄이기 위해 PCA기법으로 압축하여 고유성분의 특징을 추출하였다. 질감 특징을 추출할 수 있는 edge 히스토그램 기술자(EDH)는 통계적 접근 방법으로 edge의 공간적 분포를 나타낸다. 한 영상의 질감 특징은 16개 블록에 5개 방향성에 대한 빈도수로 총 80개의 질감 특징을 추출할 수 있다. 방향 성분은 수직, 수평, 45도, 135, 무 방향으로 총 5개이다. 각 블록은 문턱치가 넘는 방향성분에 대한 빈도수를 사용하며 한 영상에 대해 총 80개의 특징을 추출할 수 있다. 컬러모델 특징은 영상을 HMMD의 칼라공간으로 변형하여 인간의 색 지각에 매우 밀접한 특징을 가진다. 그러나 H-M-M-D 총 4개의 색 채널을 모두 특징으로 가지기엔 차원이 너무 높아서 학습, 검색, 속도에 성능 저하를 가져 올수 있다. 그러므로 본 연구에서는 차원을 축소하고 그 고유한 성질을 나타낼 수 있는 PCA기법을 이용하여 HMMD의 컬러 특징을 추출할 수 있다. 1307개 영상의 각 색 채널에 대해 PCA 기법으로 압축을 시행한다. 각 영상을 1열로 만들어 1037개 영상을 하나의 벡터로 만든다. (이 벡터는 27648x1307이 된다.) 이 벡터를 고속 PCA로 압축하면 1307x1307 기저벡터, 1307x30 고유벡터를 얻을 수 있는데 1307x30의 고유벡터는 각 1307의 영상이 한 가지 색 채널에 대해 30가지씩의 특징을 가지고 있는 것이다. 이 작업을 H-M-M-D 4가지 채널에 대해 시행하면 한 영상이 30x4의 컬러 특징을 얻을 수 있는 것이다. 또한 질감 특징까지 포함하면 한 영상이 가지는 특징은 30x4의 컬러특징과 80개의 질감 특징을 가질 수 있다. 원 영상에 대한 컬러 모델 특징 값 30x4 + 80(Edge 히스토그램 특징 값), 총 200개의 특징 값을 입력 값으로 적용하여 해마 신경망을 학습 시켰다.

표 1. 검색소요시간 비교

Table 1. The necessary time of retrieval

DB 영상개수	실험 알고리즘		
	HMMD+EDH (sec)	HMMD+EDH+PCA (sec)	HMMD+EDH+PCA+HNN (sec)
100	1043	223	5
500	1732	458	11
1000	2187	955	58
1307	2535	1135	85

※ HMMD : Hue, Max, Min, Diff

EDH : Edge Histogram

PCA : Principal Component Analysis

HNN : Hippocampal Neural Network

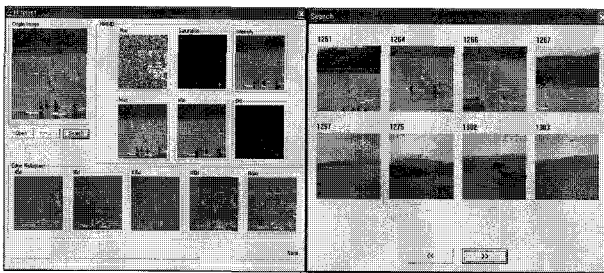


그림 6. 질의영상 1에 대한 특징 영상과 검색된 영상들.

Fig. 6. Feature values of query-1 image and retrieved result images.

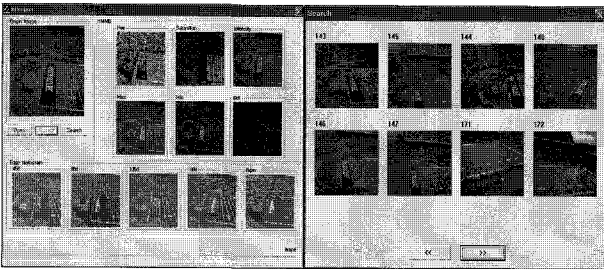


그림 7. 질의영상 2에 대한 특징 영상과 검색된 영상들.

Fig. 7. Feature values of query-2 image and retrieved result images.

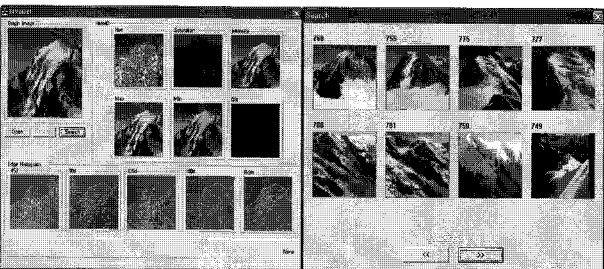


그림 8. 질의영상 3에 대한 특징 영상과 검색된 영상들.

Fig. 8. Feature values of query-3 image and Retrieved result images.

표 2. 기존 신경망과 해마 신경망과의 검색결과 비교.

Table 2. The retrieval system results.

DB 영상개수	Neocognitrom		HMM		SVM		HNN	
	Recall	Precision	Recall	Precision	Recall	Precision	Recall	Precision
100	98.2	75.7	95.4	70.4	96.3	79.9	100	93.3
500	85.6	63.2	90.5	71.7	88.7	71.9	98.3	91.1
1000	80.9	58.4	85.5	68.1	82.3	69.3	94.9	88.5
1307	73.7	52.6	79.1	60.5	76.5	63.5	90.8	82.3

※ HMM : Hidden Markov Model

SVM : Support Vector Machine

HNN : Hippocampal Neural Network

2. 실험 결과

본 연구에서는 1307개의 영상이 저장되어 있는 데이터베이스영상을 각각 질의 하였다.

표 1은 실험 알고리즘 적용하여 검색 시 소요시간을 나타내었다. HMMD 컬러모델 기술자의 특징 값들을 PCA 기법을 통해 데이터를 축소 시켰고, 해마 신경망을 통해 축소 데이터를 학습 시켰다. 표 1에서 보듯이 PCA 기법으로 데이터를 축소시키지 않고 HMMD 컬러모델 기술자와 edge 히스토그램 기술자(EDH)의 특징 값만으로 검색 시 DB 영상 개수에 따라 많은 소요시간이 걸렸다. PCA 기법과 해마 신경망을 통해 학습 시킨 후의 검색 소요시간은 상당히 줄어들었다. 검색 소요시간은 질의 영상을 100번 질의 하였을 때 검색 되는 시간을 평균적으로 걸리는 시간을 나타내었다. 그림 6, 7, 8은 검색 시스템에서 질의 영상의 특징과 검색된 영상을 보여 준다. 검색된 영상들의 번호는 DB내에 들어있는 1307개의 영상들의 번호이다. 표2는 해마 신경망으로 학습 시킨 후의 검색이 기존 신경망과 비교했을 때 보다 검색 결과가 높다는 것을 나타낸다.

V. 결론

기존의 내용기반 검색 시스템은 정확한 특징 값을 추출하기에 어려운 단점이 있었고, 실험 영상의 특징에 적합한 특징 값만을 추출함으로써 다른 여러 영상들에 대한 실험에 있어서 매우 낮은 결과를 보였다. 특징 값들에 대한 검색에 있어서 순차적인 비교를 할 경우 시간이 많이 걸리는 문제점 또한 있었다. 따라서 본 논문은 MPEG-7 국제 표준에서 제공하는 기술자들 중 컬러 질감 기술자들을 이용하여 여러 영상에 대한 동일 결과를 나타낼 수 있는 특징 값을 추출하였고, 검색의 속도를 위해 PCA 기법을 적용하여 특징 데이터를 축소 시켰다. 그리고 인간의 인지학적인 두뇌 원리인 대뇌피질과 해마 신경망을 공학적으로 모델링 하여 내용기반 특징 벡터들을 고속 학습하고, 최적의 특징을 구성할 수 있도록 해마 신경망을 적용하여 색인 및 검색 시 속도가 빠르고 정확한 검색 효과를 보였다.

본 논문에서는 1307개의 데이터베이스 영상들에 대하여 Recall과 Precision으로 평가한 결과 90.8%와 82.3%의 검색 결과를 보였고, 검색 소요 시간 또한 85초로 효과적인 검색 시스템을 구현 할 수 있었다. 향후 더 많은 실험을 통해 동영상과 연결된 검색 시스템을 구현하고자 한다.

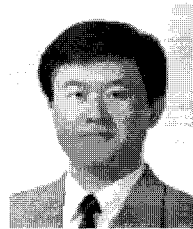
참고문헌

- [1] B. SM anjunath, P. Salembier, T. Sikora, "Introduction to MPEG-7," WILEY, 2002.
- [2] S. Antani, R. Kasturi, R. Jain, "A survey on the use of pattern recognition mehtods for abstraction, indexing and retrieval of images and video," Pattern Recognition, vol. 35, pp. 945-965, 2002.
- [3] J. D. Gibson, "Image and video processing," ACAD-EMIC PRESS pp. 687-704, 2000.
- [4] 김영호, 강대성, "인터넷에 기반한 내용기반 검색 시스템 설계," 제어·자동화·시스템공학회 논문지, 제 11권 제5호, pp. 471~475, 2005
- [5] M. Witter and H. Groenewegen, "A new look at the hippocampal connectional network," in European Neuroscience Association, 1988.



김영호

1999년 동아대학교 전자공학과 졸업. 2002년 동아대학교 전자공학과 석사. 현재 동아대학교 전자공학과 박사수료. 관심분야는 영상분할, 멀티미디어 색인 및 검색.



강대성

1984년 경북대학교 전자공학과 졸업. 1991년 Texas A&M 대학교 전자공학과 석사. 1994년 Texas A&M 대학교 전자공학과 박사. 1983년 3월~1989년 6월 국방과학연구소, 1995년 9월~현재 동아대학교 전기전자컴퓨터 공학부 부교수. 관심분야는 영상처리, 패턴인식, 영상코딩, 멀티미디어 색인 및 검색.