

투영 벡터의 형상 정보를 이용한 영상검색

(Image Retrieval Considering Shape Information of Projection Vector)

권 동 현^{*} 이 태 흥^{**}

(Donghyun Kwon) (Taihong Yi)

요약 히스토그램 인터섹션은 영상에서 컬러가 가지는 값의 빈도 수를 이용하여 간단하면서도 효율적으로 영상을 검색하는 방법으로, 영상의 글로벌 특성을 잘 나타내는 반면 영상에서의 위치 정보가 누락되어 다른 영상을 동일 영상으로 인지하기 쉽고, 영상 내에 포함된 형상 정보를 표현하는 적절한 방법은 아니다. 영상에 대한 1차원 투영을 이용하면 영상의 개략적인 형상 정보와 함께 위치 정보를 나타낼 수 있어 히스토그램의 단점을 극복할 수 있지만, 영상 크기에 따라 투영 벡터의 길이가 달라져 색인 데이터로 사용하기에는 문제가 있다. 본 논문에서는 투영벡터에서 영상이 가지는 형상 정보의 침투치를 이용하여 침투치들 간의 상대거리 및 최대 침투치에 관한 정보를 검색에 사용하였다. 검색 성능의 확인을 위하여 히스토그램 인터섹션 및 투영벡터만을 이용한 경우의 검색 결과와 비교하였고, 실험 결과를 이용하여 각 방법의 장단점을 분석하였다.

Abstract Histogram intersection method, that counts the occurrence of color pixels, is one of the easy and simple color image retrieval methods. The method has an appropriate global property but does not contain the knowledge of shape for images. The absence of spatial information makes it difficult to discriminate images of the similar histogram. The application of one-dimensional projection to each image enables to obtain shape or spatial information of image. But in this case there is another problem having different length of the projection vector according to the size of each image. Thus this paper proposes a method that uses relative distances between peaks and their maximum value in the projection vector. In order to verify retrieval performance, the experimental results between the histogram intersection method, the projection only method, and the proposed one are compared and analyzed.

1. 서론

최근 몇 년 사이 데이터 저장 장치인 디스크의 가격이 하락하고, 랩 가격도 떨어짐으로써 개인용 컴퓨터의 성능이 급속도로 향상되어 몇 년 전의 슈퍼 컴퓨터 성능을 능가하는 상황에 이르렀다. 또한 초고속 네트워크의 대중화로 인해 원하는 정보를 쉽게 검색할 수 있는 환경이 마련되었다. 그러나 기존의 검색 툴인 웹 브라우저를 통해 검색한 결과를 이용하여 원하는 정보를 찾거나

할 경우 서비스를 제공하는 웹사이트의 증가로 예전과는 달리 너무 많은 후보 결과가 검색되어 검색된 결과에서 의도하는 정보를 선별하는데 어려움이 있다. 이러한 문제 해결을 위해 향 후 검색 분야별 전문 사이트가 등장해 특정 분야에 대해 신속한 결과를 얻을 수 있도록 지원해줄 것으로 기대되고 있다.

그러나 지금까지 대부분의 검색 방법은 텍스트를 이용한 검색으로 멀티미디어 데이터에 찾고자 할 때에도 주석(annotation) 정보를 이용한 텍스트를 주로 이용하기 때문에 특정 개인이 주관적으로 만든 주석 정보에 의존하게 되어 원하는 결과를 찾기가 쉽지 않다. 따라서 멀티미디어 자체에 포함되어 있는 객관적 정보를 이용하여 특징을 추출한 후, 그 특징을 색인화 하여 검색에 이용하는 내용기반 검색기법(content based retrieval method)이 많이 연구되고 있다.

* 학생회원 : 영남대학교 전자공학과
dhkwon@popmail.com

** 비 회원 : 영남대학교 전자정보공학부 교수
tachong@ynuucc.yeungnam.ac.kr

논문접수 : 2001년 3월 26일

심사완료 : 2001년 8월 2일

인간이 일상 생활의 언어에서 사용하는 고차원적인 의미를 포함하는 의미론적 방법(semantic based method)이 실제 적용에 있어 기술적 한계성으로 인해 어려움이 많은 것과는 달리, 내용기반 검색법은 텍스트 위주의 방법과 의미론적인 방법의 중간레벨 정도를 형성하는 실용적인 방법이다.

내용에 기반을 둔 영상 검색에서는 색(color), 스케치(sketch), 형상(shape), 텍스처(texture), 공간적 관계(spatial relationship) 등의 특징을 주로 많이 이용하고 있다[1]. 그 중에서 특히 컬러는 영상의 전체 특성을 잘 나타내면서도 구현이 쉽다는 특징으로 인해 Swain[2]의 히스토그램 인터섹션을 이용한 방법이 후 활발히 연구되고 있는 분야이다. 그러나 히스토그램을 이용할 경우 공간 정보가 없어지는 단점 때문에 서로 다른 영상이라도 히스토그램만 같으면 같은 영상으로 검색하는 오류를 범하기 쉽다는 문제점이 있다.

영상의 1차원 투영(projection)은 영상에 대한 가로 및 세로 방향의 정보를 개략적으로 나타내는 방법으로, 영상에서의 위치 정보를 나타냄과 동시에 영상이 가지는 형상 정보를 나타낼 수 있는 방법이다. 그러나, 투영 기법을 사용할 경우 영상의 가로 및 세로 크기에 따라 생성되는 투영벡터의 크기가 달라지는 단점이 있어, 영상의 투영 벡터를 블록화하여 블록의 통계적 특성을 이용할 경우 영상 크기에 상관없이 검색을 위한 특징 벡터로 사용할 수 있으며, 동시에 공간정보 부재 문제를 해결할 수 있다.[3]

영상에서의 공간 정보 부재 문제 해결을 위한 방법의 일환으로 투영 기법을 사용한 접근[3]과는 달리 본 논문에서는 영상의 1차원 투영 정보를 사용하면서 투영 벡터가 가지는 영상에 대한 형상 정보로부터 검색에 필요한 특징을 추출하여 검색에 이용하였다. 투영 벡터에서 변화가 심한 부분을 나타내는 첨두치를 추출한 후, 색인 데이터 생성을 위하여 첨두치 간의 상대거리 및 첨두치 중 최고값에 관한 정보를 이용하여 영상의 형상 정보가 색인 데이터에 포함되도록 하였다.

본 논문의 2절에서는 투영기법에 대한 소개를 하였으며, 3절에서는 본 논문에서 제안하는 방법을 데이터를 이용해 분석하면서 제시하였고, 4절에서는 제안된 방법에 대한 실험 및 결과를 분석과 함께 나타내었고, 5절에서 결론을 맺었다.

2. 투영을 이용한 영상검색

컬러를 이용한 검색 중 최초로 소개된 히스토그램 인터섹션은 질의 영상에 대한 히스토그램을 Q 라 하고, 데

이타베이스 영상에 대한 히스토그램을 I 라 할 때

$$HI = \sum_{i=1}^N \min(I_i, Q_i) \quad (1)$$

와 같이 질의 영상에 대해 데이터베이스 영상에서 구한 히스토그램과의 차이 값을 합산하여 비슷한 정도를 표현하고, 그 차이가 가장 적은 영상을 가장 유사한 영상으로 검색하였다. 그러나 서로 다른 영상이라도 히스토그램은 같을 수 있기 때문에, 영상 내에서의 컬러 빈도 수만을 비교 대상으로 하는 컬러 히스토그램을 이용할 경우 잘못된 영상을 검출할 가능성이 많다. 이러한 히스토그램의 단점을 보강하기 위해 Stricker[4]는 영상을 임의의 영역으로 나눈 후 모멘트를 계산하여 색인 데이터로 이용하였고, Smith[5]는 히스토그램 back-projection 기법을 이용하여 영상을 분할 후 이진 영상화 한 정보에서 색인 정보를 추출한 후 검색에 이용하였으며, Pass[6]는 영상 정보를 응집 정도에 따라 CCV(Color Coherence Vector)로 분류하여 색인 정보를 추출하였다. 본 논문에서는 이들과는 달리 투영 기법을 이용하여 새로운 방법을 제안하려 한다.

먼저 이진 영상에서의 투영은 그림 1과 같이 나타낼 수 있다.

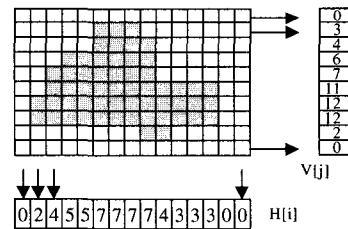


그림 1 이진 영상에서의 투영 예

크기가 $N \times N$ 인 영상을 $I[r,c]$ 라 할 때 수평 및 수직 방향으로의 투영은 각각

$$H[r] = \sum_{c=0}^{N-1} I(r,c) \quad (2)$$

$$V[c] = \sum_{r=0}^{N-1} I(r,c) \quad (3)$$

와 같이 나타낼 수 있다.

영상에서의 1차 모멘트를 이용하여 물체의 위치를 알 수 있는 것까진, 영상에서의 1차 모멘트가 투영에서의 1차 모멘트로 나타난다는 성질을 이용하여 투영된 벡터에서 물체의 위치를 계산할 수 있다. 식

$$A = \sum_{c=0}^{N-1} V[c] = \sum_{r=0}^{N-1} H[r] \quad (4)$$

을 이용하여 수평 및 수직 방향의 위치를 계산하면 식

$$\bar{y} = \frac{\sum_{r=0}^{N-1} rH[r]}{A} \quad (5)$$

$$\bar{y} = \frac{\sum_{c=0}^{N-1} cV[c]}{A} \quad (6)$$

와 같이 행방향 y 및 열방향 x 의 위치를 계산할 수 있다[7]. 그러나 투영은 영상의 크기에 따라 생성되는 행 및 열 방향의 투영 벡터 크기가 달라진다는 단점이 있다. 이러한 투영 벡터의 단점을 극복하고자 투영 벡터의 행 및 열 방향에 대해 블록을 구성한 후, 각 블록에 대해 각각 투영 벡터의 통계적 성질을 적용하면 적은 양의 색인 데이터로 영상 검색을 가능하게 할 수 있다.[3]

3. 투영 벡터의 침두치 간 거리를 이용한 영상검색

영상 검색을 위해서는 영상 데이터가 가지고 있는 값들의 분포를 이용하여 영상 내의 물체의 형상을 나타낼 수 있는 적절한 방법을 찾아야 한다. 그러나 히스토그램은 영상 전체의 특성은 훌륭하게 표현할 수 있지만 각 컬러가 가지는 값의 개수만을 계산하기 때문에 영상 내의 물체에 대한 형상 정보를 표현하기에 적합한 방법은 아니다. 이러한 단점을 극복하기 위한 기존의 방법[4][5][6]에서 직간접적으로 영상 분할을 통해 색인 정보를 구하는 시도를 하였고, 패턴 인식분야에서 형상을 표현하기 위한 일반적인 방법으로 영상 분할(segmentation)이 많이 사용되지만 본 논문에서는 영상 분할 없이 영상 내에 포함된 물체를 나타내는 특징을 추출하여 검색에 이용하기 위하여 투영 기법을 사용한다.

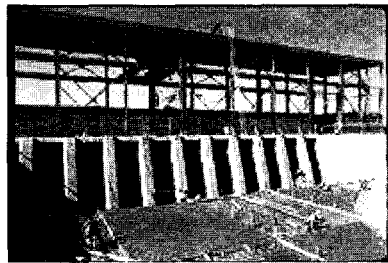
영상을 1차원으로 투영할 경우 영상 내의 각 물체가 존재하는 윤곽 영역에서 투영 벡터 값에 변화가 일어날 것이고, 물체 내부나 배경 등의 평탄한 영역에서는 투영 벡터의 값도 평이하게 표현될 것이다.

그림 2는 예제 영상에 대한 수평 및 수직 방향으로의 투영을 나타낸 것으로 (a)의 예제 영상에 대해 (b),(c)의 투영 데이터가 영상의 형상 정보를 개략적으로 나타내면서 영상 내에서의 윤곽 부분에서 투영 벡터 값이 급격하게 변화하고 있다는 것을 나타내고 있다.

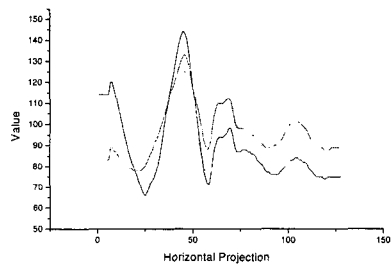
본 논문에서는 투영벡터에서 특징을 추출하기 위하여 영상에 대해 구한 1차원 투영 벡터에서 각 방향의 형상 정보를 나타내는 봉우리에서 제일 높은 값을 나타내는 침두치들을 검출해 낸 후, 각 침두치 간의 거리를 이용하여 검색을 위한 데이터로 이용하고자 한다. 이 경우 각 침두치들간의 거리 정보는 물체들이 가지는 형상 정보의 봉우리들의 위치를 대표적으로 표현한다고 볼 수

있다.

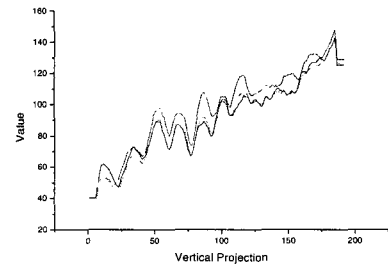
또한 영상 내에서 질의 영상과 데이터베이스 영상 간에 물체의 이동이 있더라도 검색 결과에 영향을 주지 않기 위해서는 상대적인 개념을 적용한 특징을 이용하는 것이 바람직하기 때문에 검색을 위한 색인 데이터는 추출한 침두치들 중 최대 값을 가지는 것에서부터 정렬(sort)을 하여 최대 값을 포함한 4개까지를 선택한 후, 각 침두치 간의 상대거리를 이용하였다. 실제 적용을 위해서는 최대 침두치의 위치 또는 최대 침두값을 함께 색인 데이터로 이용하여 영상 데이터가 가지는 특성을 반영하고자 하였다. 따라서 색인 데이터는 세 개의 각 컬러 채널 당 수직과 수평으로 각각 1개의 최대 침두치에 해당하는 데이터와 3개의 상대거리 데이터가 저장되어 하나의 영상 당 24개의 데이터가 색인 데이터로 사용되었다.



(a) 예제 영상



(b) 수평 방향으로의 투영벡터



(c) 수직 방향으로의 투영 벡터

그림 2 예제 영상에 대한 투영 벡터

4. 실험 및 실험 결과

제한된 방법의 성능 확인을 위하여 영상 데이터베이스에 대해 히스토그램 인터섹션을 이용한 방법, 투영만을 이용한 방법, 투영 벡터의 침두치간 상대거리를 이용한 방법으로 나누어 적용한 후 그 검색 결과를 비교하였다.

실험에 사용된 영상 데이터베이스는 Chabot 시스템[8]에서 사용한 영상 데이터베이스 중 177개를 기용하였는데 여러 환경에서 촬영한 다양한 형태의 영상을 포함하는 192 x 128 크기의 영상들이다. 질의 영상으로는 177개의 영상 중 30개를 선정하여 각 질의 영상에 대한 유사 영상을 미리 정해 놓은 후, 각 질의 영상을 query by example 형식으로 주었을 때 질의 영상과 유사한 영상이 검색되는 순위를 기록하였다.

히스토그램 인터섹션, 투영만을 이용한 방법의 경우 영상을 RGB 채널로 분리한 후, 각 채널에 대해 히스토그램 및 투영 데이터를 구하여 색인 데이터로 저장하였으며, 투영 벡터의 상대 거리를 이용하는 경우 영상의 RGB 채널별로 구한 투영 벡터에 대해 원도 크기 11의 1차원 지역 통과 필터를 적용하여 투영 벡터의 윤곽을 스무딩 하였다. 스무딩한 데이터에 대해 7준 위치를 중심으로 거리 4이상 상승한 후, 거리 6이상 하강하는 점을 침두치로 취하여, 각 침두치들 중 상위로부터 네번째까지의 값을 가지는 위치들 간의 상대 거리를 구하여 검색을 위한 색인 데이터로 저장하였다. 색인 데이터로는 이들 침두치간 상대거리와 더불어 최대 침두치의 위치를 저장한 경우와, 최대 침두치의 값을 저장한 경우 두가지로 나누어 실험하였다.

질의 영상에 대한 검색 결과를 산출하기 위한 유사도 비교 척도(similarity measure)로는 질의 영상을 Q라 하고 데이터베이스 영상을 I라 할 때

$$S_{ij} = \sum_{k=1}^4 \sum_{l=1}^4 |I_{ij} - Q_{kl}| \quad (7)$$

와 같이 두 차이 값의 절대치를 사용하였다.

표 1에서는 히스토그램 인터섹션을 이용하였을 때와 투영 데이터만을 이용한 경우, 투영 벡터의 침두치 상대 거리에 최대 값의 위치를 이용한 경우 및 최대 값을 이용한 경우에 대한 검색 결과를 나타낸 것이다.

표 1의 결과에서 확인할 수 있는 바와 같이 5위 이내의 검색율을 보면 투영 벡터의 침두치간 상대거리와 침두치의 최대값을 같이 사용한 경우가 가장 높은 검색 성공율을 보이고 있고, 10위 이내에서는 대부분의 방법들이 거의 비슷한 양상의 검색율을 보인 반면, 15위 이

내에서의 검색율에서는 투영 벡터의 침두치 상대거리 및 최대값의 위치 정보를 이용한 경우의 방법이 가장 좋은 검색 결과를 보여주고 있다는 것을 알 수 있다.

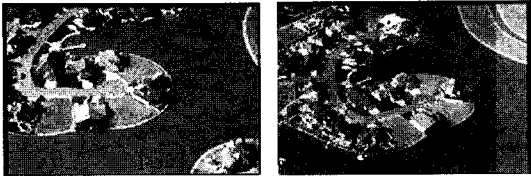
표 1 검색 결과

검색 방법	5위 이내	10위 이내	15위 이내
HI	66.7%	80.0%	83.3%
projection	73.3%	86.7%	93.3%
Prj+max pos.	60.0%	80.0%	96.7%
Prj+max val.	76.7%	80.0%	83.3%

이 결과를 통해 5위 이내에서는 질의 영상과 비슷한 영상의 형상 정보를 데이터베이스 영상에서 제대로 검색하기 위해서는 투영벡터와 그 형상 정보를 특징으로 삼는 형상의 침두 최대치가 큰 역할을 한다는 것을 알 수 있으며, 15위 이내에서는 값 자체보다는 위치가 더 큰 역할을 한다는 것을 알 수 있다. 즉 정확한 검색을 위한다면 값 정보를 더 강조할 수 있고, 특정 범위 내에서 선택할 의향이면 위치 정보를 더 강조할 수 있다. 또한 전체적 측면에서 침두값의 최대값 위치 및 상대거리를 이용한 경우가 최대값 및 상대거리를 이용한 경우보다 결과가 좋다는 것을 알 수 있는데, 이는 질의 영상과 유사한 데이터베이스 내의 영상들이 포함하는 물체의 위치가 크게 변화하지는 않았다는 것을 말해주는 것이다. 만약 데이터베이스 내의 영상들이 큰 위치 변동이 있을 경우 침두치에 상대 거리를 이용한 경우가 더 좋은 결과를 나타내리라는 것을 기대할 수 있다.

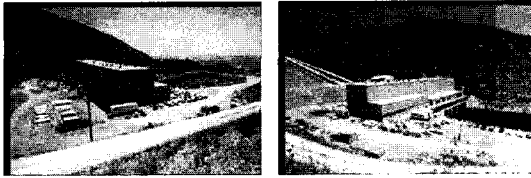
검색 내용을 분석해 볼 때 그림 3 (a)의 질의 영상이 주어질 때 (b)의 영상을 찾는 경우 히스토그램 인터섹션의 경우는 15위 이내에서 검색을 해 내지 못한 반면 나머지 세 경우에 대해서는 1위의 성공율을 보여 히스토그램의 단점이 가장 잘 드러나면서 투영을 사용한 경우의 형상 정보가 가장 효과를 발휘한 경우로 볼 수 있으며, 그림 4의 경우 (a)영상을 질의 영상으로 주었을 때, (b)의 영상을 히스토그램 인터섹션에서는 5위로 검색해 낸 반면, 투영만을 이용한 경우와 투영 벡터의 값 및 상대거리를 이용한 경우에는 15위 이내에서 검색해 내지 못했고, 투영 벡터의 최대치의 위치 및 상대거리를 이용한 경우에는 12위로 검색해 내었다. 이 경우는 컬러 값의 빈도만으로는 두 영상이 유사하지만 영상 정보 측면에서는 두 영상을 다른 영상으로 밖에 볼 수 없다는 측면에 기인한다고 볼 수 있다.

또한 그림 5와 같은 경우는 (a)의 질의 영상에 대한



(a) 질의 영상 (b) 유사영상

그림 3 분석 예제 영상 1



(a) 질의 영상 (b) 유사 영상

그림 4 분석 예제 영상 2



(a) 질의 영상 (b) 유사 영상

그림 5 분석 예제 영상 3

(b)의 검출 결과를 보면 히스토그램 인터섹션의 경우 8위를 보인 반면 투영만을 사용한 경우는 2위로 검출해 내었고, 최대치 위치 및 상대거리를 이용한 경우에는 7위의 결과를 나타낸 반면 최대값 및 상대거리를 이용한 경우에는 1위로 검출해 내어 히스토그램 등의 컬러 빈도나 자세한 형상 정보를 이용하는 경우 검색이 어려운 데이터베이스의 검색에 형상 정보를 나타내는 투영 벡터에서 특징을 추출하면 좋은 결과를 기대할 수 있다는 것을 알 수 있다.

색인 데이터량을 비교할 경우 각 영상에 대해 히스토그램은 256 x 3개이고, 투영만을 사용한 경우는 192 + 128개인 반면 제안된 방법의 경우는 24개만을 사용하면 되기 때문에 대용량 데이터베이스의 경우에는 제안된 방법이 빠른 응답을 나타낼 있다는 것을 알 수 있다.

5. 결론

본 논문은 기존의 컬러 히스토그램이 영상 내에서 컬러의 빈도를 이용하여 영상을 검색하는 방법과는 달리 영상의 개략적인 형상 정보를 이용하는 투영을 적용하

여 영상 검색을 수행하는 방법에 관한 논문이다. 본 논문에서 사용하는 방법은 영상에서 1차원 투영 벡터를 구한 후, 각 투영 벡터가 가지는 형상 정보의 침두치를 구하여 정렬한 후, 각 침두치 들 간의 상대거리를 색인 데이터로 이용하여 검색을 수행하고 있다.

이 때의 침두치간 거리는 영상의 형상 정보를 개략적으로 표현하는 것으로, 제안된 방법을 컬러 히스토그램 인터섹션 방법, 투영만을 사용한 검색법 등과 비교하여 실험을 한 결과, 검색율에서 5위 이내에서는 제안된 방법이 약 10% 정도로 우수성을 나타내었으며, 검색된 데이터들 통하여 각 방법들 간의 장단점을 분석한 결과 영상 내의 색채 성분이 거의 변화하지 않는 환경에서의 검색에서는 컬러 히스토그램이 우수하나, 같은 형상이면서도 촬영 각도에 따라 컬러 차이가 나는 영상인 경우의 검색에는 제안된 방법의 검색 결과가 우수하게 나타난다는 결론을 내릴 수 있었다.

참 고 문 헌

- [1] F.Idris and S.Panchanathan, "Review of Image and Video Indexing Techniques," *Journal of Visual Communication and Image Representation*, vol.8, no.2, pp.146-166, June 1997.
- [2] Michael J. Swain and Dana H. Ballard. "Color indexing," *International Journal of Computer Vision*, vol.7, no.1,pp. 11-32, 1991.
- [3] 권동현 외 3인, "투영벡터의 통계적 성질을 이용한 영상검색", *한국통신학회 논문지* 제 25권 제7A호, pp. 1044~1049, Jul. 2000.
- [4] Markus Stricker and Alexander Dimai, "Color Indexing with Weak Spatial Constraint," *Storage and Retrieval for Image and Video Databases IV, SPIE proceedings*, vol. 2670, pp. 29-40, 1996.
- [5] John R. Smith and Shih-Fu Chang, "Tools and Techniques for Color Image Retrieval," *SPIE proceedings*, vol. 2670, pp. 1630-1639, 1996.
- [6] Greg Pass and Ramin Zabih, "Histogram Refinement for Content Based Image Retrieval," *IEEE Workshop on Applications of Computer Vision*, pp. 96-102, 1996.
- [7] Ramesh Jain, *Machine Vision*, pp.35-38, McGraw-Hill 1995.
- [8] Virginia E. Ogle and Michael Stonebraker, "Chabot : Retrieval from a Relational Database of Images," *IEEE Computer*, vol. 28, no. 9, 1995.



권 동 현

1990년 2월 영남대학교 전기공학과 졸업.
 1990년 8월 ~ 1992년 8월 영남대학교
 전자공학과 석사 졸업. 1992년 8월 ~
 1998년 6월 한국전자통신연구원 연구원.
 1998년 9월 ~ 현재 영남대학교 전자공
 학과 박사과정. 관심분야는 영상 및 비디

오 압축/신호처리, 영상/비디오 검색, HDTV 등



이 태 홍

1975년 2월 서울대학교 공업교육학과 전
 자공학전공 졸업. 1976년 5월 ~ 1976년
 10월 동양정밀(주) 중앙연구소 책임연구원.
 1980년 2월 서울대학교 대학원 공업교육
 학과 전자공학전공 졸업(석사) 1980년 1
 월 ~ 1981년 2월 금성전기(주)기술연구소
 기좌. 1981년 3월 ~ 현재 영남대학교 전자정보공학부 교
 수. 1990년 미국 Syracuse University 교환교수. 1996년
 10월 ~ 1997년 10월 한국통신학회 대구·경북 지부장.
 1998년 미국 Univ. of Washington 교환교수. 1997년 2월
 서울대학교 대학원 전자공학과 졸업(박사). 관심분야는 신
 호처리, 영상/비디오 처리