

# 영상편집효과를 고려한 내용기반 영상 검색의 개선에 관한 연구

강석준<sup>†</sup>, 배태면<sup>\*\*</sup>, 김기현<sup>\*\*\*</sup>, 노용만<sup>\*\*\*\*</sup>, 한승완<sup>\*\*\*\*\*</sup>, 정치운<sup>\*\*\*\*\*</sup>, 남택용<sup>\*\*\*\*\*</sup>

## 요 약

멀티미디어 콘텐츠가 급격히 증가함에 따라 사용자들은 다양한 유통 경로를 통하여 많은 멀티미디어 콘텐츠를 이용할 수 있게 되었다. 내용기반 영상 검색시스템은 영상 데이터의 내용을 다양한 시각적 특징 값들로 표현하여, 수많은 영상 중에서 사용자가 원하는 영상을 검색하고 원하지 않는 영상을 필터링 하도록 한다. 그러나 멀티미디어 데이터의 편집은 영상 데이터의 고유한 시각적 특징 값들을 왜곡시켜 잘못된 검색 결과나 필터링 결과를 제공하여 내용기반 영상 검색시스템의 성능을 저하시킨다. 본 논문에서는 이러한 영상편집효과가 가운데 글자삽입, 프레임의 삽입, 그리고 여러 영상으로의 구성과 같은 편집효과에 대해 분석하고 이러한 편집효과를 제거하는 알고리즘을 고려한 내용기반 검색시스템을 제안하였으며, 실험을 통해 향상된 검색 결과를 확인할 수 있었다.

## Improvement of Content-based Image Retrieval by Considering Image Editing Effect

Seok Jun Kang<sup>†</sup>, Tae Meon Bae<sup>\*\*</sup>, Ki Hyun Kim<sup>\*\*\*</sup>, Yong Man Ro<sup>\*\*\*\*</sup>,  
Seung Wan Han<sup>\*\*\*\*\*</sup>, Chi Yoon Jeong<sup>\*\*\*\*\*</sup>, Taek Yong Nam<sup>\*\*\*\*\*</sup>

## ABSTRACT

With the rapid increase of the number of multimedia contents, people consume a lot of multimedia contents through various distribution channels. Content-based image retrieval uses visual features that represent the contents of images. And users can retrieve or filter images based on the contents of the images using the features. But, the editing of the multimedia contents distorts the original visual features of the multimedia contents, thereby the performance of content-based image retrieval system could be lowered. In this paper, we describe the image editing effects that lower the performance of the retrieval system and propose algorithms that can remove the image editing effect and improve content-based image retrieval system.

**Key words:** Content-based image Retrieval System(내용기반 영상검색 시스템), Image Editing(영상편집), Image Retrieval(영상검색), Image Filtering(영상 필터링)

※ 교신저자(Corresponding Author) : 노용만, 주소 : 대전광역시 유성구 문지동 한국정보통신대학교(305-732), 전화 : 042)866-6289, FAX : 042)866-6245, E-mail : yro@icu.ac.kr  
접수일 : 2005년 9월 28일, 완료일 : 2006년 2월 17일

<sup>†</sup> 한국정보통신대학교  
(E-mail : impedance99@hotmail.com)

<sup>\*\*</sup> 준회원, 한국정보통신대학교  
(E-mail : heartles@icu.ac.kr)

<sup>\*\*\*</sup> 정회원, 한국정보통신대학교

(E-mail : kihyun@icu.ac.kr)

<sup>\*\*\*\*</sup> 종신회원, 한국정보통신대학교 교수

<sup>\*\*\*\*\*</sup> 정회원, 한국전자통신연구원 정보보호연구단  
(E-mail : hansw@etri.re.kr)

<sup>\*\*\*\*\*</sup> 정회원, 한국전자통신연구원 정보보호연구단  
(E-mail : iamready@etri.re.kr)

<sup>\*\*\*\*\*</sup> 정회원, 한국전자통신연구원 정보보호연구단  
(E-mail : tynam@etri.re.kr)

## 1. 서 론

멀티미디어 콘텐츠의 증가와 함께, 웹 또는 E-mail 상에 존재하는 수많은 멀티미디어 콘텐츠 중에서 사용자가 원하는 콘텐츠를 검색하거나 원하지 않는 정보를 필터링하는 기능이 요구되어지고 있다. 멀티미디어 콘텐츠를 구성하는 영상 및 음성데이터의 내용은 여러가지 영상 및 음성 특징정보들의 조합으로 표현 될 수 있고[1-4], 이를 이용한 내용기반 검색에 대한 많은 연구가 진행되어왔다[5-8]. 또한, 내용기반 영상 검색 및 필터링 시스템은 자연스러운 사용자 질의를 통해 효과적이고 효율적으로 멀티미디어 데이터를 검색할 수 있다는 장점으로 인해 다양한 분야에서 점차 많이 활용되고 있다.

내용기반 영상 검색시스템은 영상의 다양한 시각적 특징 정보들을 이용한다. 칼라, 질감, 형태와 같은 영상 특징정보가 대표적인 것들로 이러한 특징정보를 수치화하고 조합하여 영상의 내용을 표현한다. 그러나 멀티미디어 편집도구의 보급으로 인하여 웹 또는 E-mail 상에 존재하는 대부분의 영상들은 수정 및 편집되어, 영상특징정보는 원본 영상과는 많이 다른 특징값을 가지게 되고 결과적으로 내용기반 검색의 성능을 저하시킨다. 특히, 영상에 삽입된 글자, 영상 외곽에 삽입된 프레임, 그리고 여러 영상으로 구성된 영상은 영상의 특징을 심하게 변질시킨다. 이러한 영상 편집으로 인한 잘못된 검색 또는 필터링을 방지하기 위해 영상 편집효과를 고려한 새로운 접근 방법을 제안한다.

본 논문에서는 영상 편집에 의한 내용기반 검색시스템의 성능저하를 실험을 통해 확인 및 분석하고

검색 성능 향상을 위해 전처리 과정을 통하여 편집효과를 제거한다. 또한 편집효과를 고려한 내용기반 검색시스템을 제안하여 기존의 내용기반 검색시스템의 성능을 향상시킬 수 있음을 확인하였다. 먼저, 본 논문의 2장에서는 내용기반 검색시스템에서 사용되는 대표적인 특징정보와 영상 편집에 의한 특징정보들의 왜곡을 확인 및 분석을 하였다. 그리고 3장에서 특징정보 왜곡을 방지하기 위한 전처리 알고리즘과 전처리 과정이 포함된 내용기반 검색시스템을 제안하였다. 마지막으로 실험을 통해 제안한 내용기반 검색시스템의 성능향상을 확인하고, 5장에서 결론을 맺는다.

## 2. 내용기반 영상 검색시스템과 영상 편집효과

### 2.1 내용기반 영상 검색시스템

멀티미디어 콘텐츠는 정지영상, 동영상, 오디오 등 다양한 형태의 데이터들로 구성되어 있고, 내용기반 검색시스템은 각 데이터의 특징정보들을 수치화하여 내용을 표현한다. 정지영상에 대한 일반적인 내용기반 검색 기법들은 영상의 칼라나 질감, 형태 등의 특징정보를 수치화하여 내용을 표현한다. 먼저 데이터베이스에 저장된 영상의 특징정보들을 자동으로 추출하고 영상과 함께 인덱싱되어 데이터베이스에 저장된다. 그 후 사용자는 원하는 특징정보를 포함하는 영상을 요청하거나 영상의 모양을 스케치하는 방식으로 내용기반 시스템에 질의를 할 수 있다[5]. 마지막으로 내용기반 영상 검색시스템은 사용자 질의 영상에서 추출한 특징정보와 가장 유사한 영상들을 반환한다. (그림 1)은 이러한 과정을 보여주고 있다.

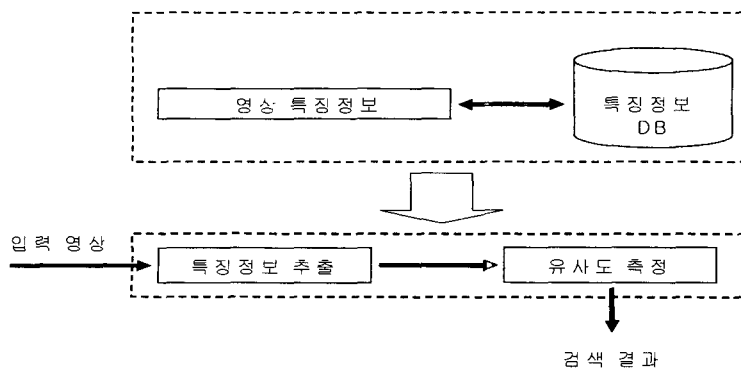


그림 1. 내용기반 영상 검색시스템 개요도

현재까지의 내용기반 영상 검색 기법에 대한 연구는 검색의 정확성을 향상시키거나 검색에 필요한 연산량을 최소화하는데에 중점을 두고 연구되어 왔다. 정확하고 빠른 영상 검색을 위해, [6]에서는 영상 데이터로부터 형태적 특징정보와 색상 특징정보를 추출하여 원하는 영상을 검색할 수 있도록 하였다. 칼라 히스토그램과 형태 정보를 추출하여 영상의 비교에 사용하였으며, 정확하고 빠르게 비교 검색하기 위해서 색상 성분에 대한 특징정보를 트리로 만들고, 형태관련 특징정보를 변형된 트리로 인덱싱하는 방법을 제안하였다. 특징정보를 포함하는 일반적인 질의 영상뿐만 아니라, 스케치된 영상을 입력받는 대표적인 시스템은 IBM 연구소에서 개발된 QBIC이다 [7]. 이 시스템은 사용자에게 영상의 색상이나 질감, 모양 등과 같은 다양한 속성에 기반하여 시각적 질의를 제공할 뿐만 아니라 제한된 범위내의 비디오 검색이 가능하다. QBIC은 또한 사용자로 하여금 데이터베이스내의 영상에 대하여 적절한 키워드로 주석을 붙일 수 있도록 허용함으로써, 비록 제한적이지만 주석에 기반한 영상 검색도 가능하다. 반면에 VisualSEEK은 웹에 기반하여 원하는 영상을 검색하는 내용기반 영상 검색시스템으로써 QBIC에서와 유사한 검색을 제공할 뿐만 아니라 영상 내의 객체영역에 대한 색상과 객체영역의 공간정보 사이의 상관관계를 더 자세히 묘사할 수 있는 인터페이스를 제공한다[8]. 이 시스템은 객체영역의 크기와 공간적 위치정보에 대한 향상된 인덱싱 기법을 이용함으로써 복잡하게 결합된 다양한 색상 및 공간 그리고 색상 및 질감에 기반한 영상 질의를 제공한다. 정지영상에 대한 내용기반 검색의 다른 방법은 영상의 가공되지 않은 원본 데이터(raw data)로부터 직접 특징을 추출하는 것이 아니라 일단 직교변환 등을 시켜 변환된 데이터로부터 특징들을 추출하는 방법이다. [9]에서는 영상의 색, 질감, 형태의 모든 특징을 함께 웨이블릿 변환에 적용하였다. 또한 영상의 웨이블릿 계수를 통한 내용 표현 문제 뿐 아니라 데이터 압축 문제도 하나의 통합된 프레임워크에서 연구하였다.

내용기반 영상 검색 기법은 특징정보의 선택, 추출, 인덱싱, 그리고 비교하는 알고리즘을 통해 사용자의 질의 영상과 비슷한 영상을 신속하고 정확하게 검색하는 것인데, 이것은 왜곡이 없는 환경에서 영상을 검색한다고 볼 수 있다. 그러나 웹이나 E-mail

사용 환경의 특성상 제공되는 멀티미디어 콘텐츠는 다양한 편집효과가 적용되고, 이러한 편집효과는 영상특징정보를 많이 왜곡시킨다. 특히, 영상에 삽입된 글자, 영상의 외곽에 꾸며진 프레임, 그리고 내용이 다른 부분 영상의 삽입과 같은 편집은 영상특징정보들을 가장 심하게 왜곡시키는 편집으로써 내용기반 영상 검색시스템의 성능저하의 요인이 된다.

## 2.2 영상 편집에 따른 검색 성능 및 편집효과 분석

글자가 삽입된 영상, 프레임이 삽입된 영상, 그리고 여러 영상으로 구성된 영상은 웹이나 E-mail을 통해서 많이 볼 수 있는 편집효과들으로써, 영상특징정보가 왜곡되어 검색이 어려운 경우이다.

먼저, 글자가 삽입된 영상의 경우, 영상에 삽입된 글자가 원본영상의 질감을 왜곡시키는 것을 확인할 수 있다. 영상에서 객체는 중앙에 위치하기 때문에 글자는 대부분 외곽에 삽입되며, 글자의 크기나 형태는 다양하다. 일반적으로 글자가 크면 전체 영상의 칼라정보를 왜곡시키며, 글자가 작고 촘촘하면 질감을 왜곡시킨다. 프레임이 삽입된 영상의 경우, 프레임은 영상을 꾸미기위해 일반적으로 영상의 외곽에 위치한다. 프레임의 형태는 사각형이나 원 또는 기하학적인 모양을 가지지만 칼라변화가 거의 없고 프레임의 질감 또한 균일하다. 그러나 삽입된 프레임의 칼라와 질감이 원본 영상에 혼합되어, 프레임이 삽입된 영상은 원본 영상과는 다른 칼라정보와 질감정보를 가지게 된다. 마지막으로 여러 영상으로 구성된 영상의 경우, 여러영상을 구성하는 각각의 영상이 서로 다른 내용을 가지고 있어서, 여러영상 중에서 사용자 질의와 유사한 영상이 있지만, 사용자 질의 영상과 특징값이 다른 나머지 영상들이 혼합되어 전체영상을 구성하기 때문에 검색 또는 필터링시에 정확한 결과를 얻기 힘들다. (그림 2), (그림 3), (그림 4)는 글자가 삽입된 영상, 프레임이 삽입된 영상, 그리고 여러영상으로 구성된 영상의 형태와 실례를 보여준다.

본 논문에서는 실험을 통하여 편집된 영상에 대한 내용기반 검색시스템의 성능을 살펴보고, 편집효과를 시각적으로 제거한 후에 내용기반 필터링시스템의 성능을 살펴보았다. 실험에서 사용된 내용기반 필터링시스템은 확률에 기반한 베이시안(Bayesian) 분류기를 사용하였으며, 베이시안 분류기의 학습에

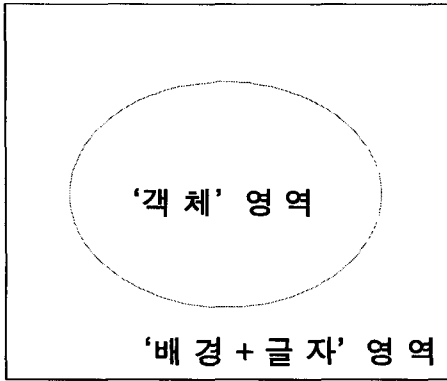


그림 2. 글자가 삽입된 영상의 형태 및 예

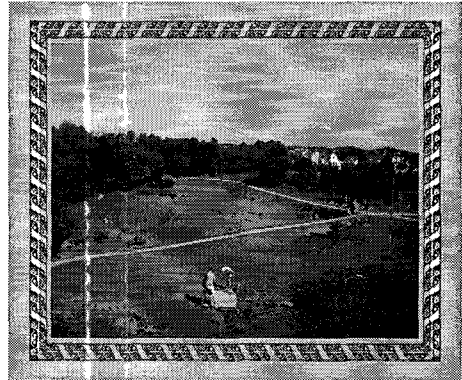
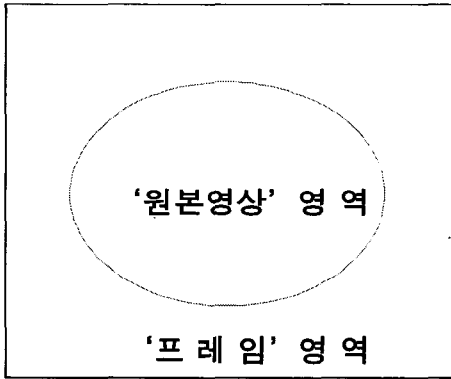


그림 3. 프레임이 삽입된 영상의 형태 및 예

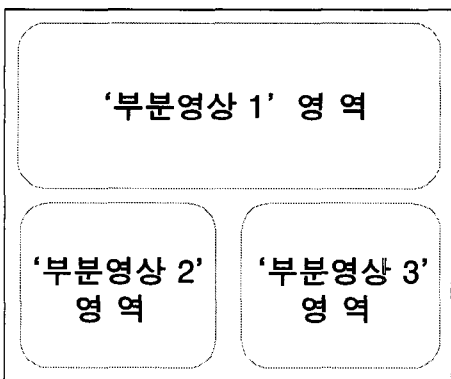


그림 4. 여러영상으로 구성된 영상의 형태 및 예

사용된 영상은 검출 성능 평가시 제외하였다. 베이시안 분류기는 두 종류의 영상(유해영상, 비유해영상)을 MPEG-7에서 정의하는 칼라히스토그램, 에지, 질감정보를 이용하여 학습하여 유해영상을 검출하도록 하였다. (표 1)은 편집되지 않은 영상과 편집된

영상들에 대한 내용기반 필터링시스템의 성능을 보여준다. 편집된 영상에 대한 검출률에서 볼 수 있듯이, 글자가 삽입된 영상, 프레임이 삽입된 영상, 그리고 하나의 영상이 각기 다른 내용을 가지는 여러영상으로 구성된 영상에 대해서는 편집되지 않은 영상보

표 1. 편집유형에 따른 내용기반 필터링시스템의 검출률

편집 유형	총 개수	검 색 률
편집되지 않은 영상	688	0.75
글자가 삽입된 영상	149	0.13
프레임이 삽입된 영상	112	0.33
여러영상으로 구성된 영상	103	0.52

표 2. 편집효과를 제거하여 재구성된 영상들에 대한 검출률

편집된 영상	총 개수	검 색 률
글자가 삽입된 영상	149	0.64
프레임이 삽입된 영상	112	0.62
여러영상으로 구성된 영상	103	0.62

다 낮은 검출률을 보이는 것을 알 수 있다.

편집효과에 의해 검출률이 낮아진다는 것이 올바른 가정이려면, 편집효과를 제거함으로써 검출률의 향상을 예상할 수 있다. 실험을 통해, 다음과 같은 규칙에 따라 편집효과를 시각적으로 제거하여 영상을 직접 재구성 하였다. 먼저 글자가 삽입된 영상은 전체영상에서 글자가 삽입된 영역을 사각영역으로 결정하고, 그 사각영역을 흰색으로 채워 영상을 재구성한다. 프레임이 삽입된 영상은 전체영상에서 프레임을 제거하기위해 프레임을 제외한 가장 큰 사각형의 부분영상을 추출한다. 마지막으로 여러영상으로 구성된 영상은 각각의 부분영상들로 분할하여 각 부분영상을 하나의 독립적인 영상으로 정의하였다. (표 2)에서 볼 수 있듯이, 공간영역에서 편집효과를 제거하여 재구성한 경우, 편집효과를 제거하기전과 비교해서 검출률이 향상되는 것을 확인할 수 있으며, 이러한 방법으로 영상의 편집효과를 효과적으로 제거한다면 검색 및 필터링 시스템의 성능향상이 가능함을 알 수 있다.

### 3. 편집효과를 고려한 내용기반 영상 검색시스템

#### 3.1 글자가 삽입된 영상의 편집효과 제거

영상에서 글자가 삽입된 영역을 검출하는 방법에 대해서는 많은 연구가 이루어져 왔다[10-15]. 그 가운데 연결요소를 사용하는 방법[14]과 질감정보를 이용하는 방법[15]이 대표적이다. 현재의 글자영역의 검출 알고리즘들은 대부분 높은 연산량이 필요하

다. 그러나 본 논문에서는 글자 삽입여부 및 글자영역의 검출이 내용기반영상 검색의 전처리로 적용되므로, 글자 삽입 영역 검출의 높은 연산량을 피하기 위해 글자가 삽입된 영상여부를 먼저 판별하고 글자가 삽입되었다고 판단되는 영상에 대하여 글자가 삽입된 영역을 검출하도록 하였다.

(그림 5)는 일반 영상과 글자가 삽입된 영상을 포함한 500장의 영상에서 구한 MPEG-7에서 정의하고 있는 균질질감(Homogeneous Texture) 값의 평균과 편차를 나타내고 있다. (그림 5)와 같이 글자가 삽입된 영상의 질감정보는 글자가 삽입되지 않은 영상과 특성이 다르게 나타나는 것을 알 수 있다. 특히 (그림 5)에서의 1~6번째 특징값이 크게 차이가 남을 알 수 있다. 따라서, '균질질감'의 1~6번째 특징정보를 사용한다면 글자가 삽입된 영상의 통계적인 검출이 가능함을 알 수 있다.

(그림 6)은 입력된 영상의 글자 삽입 여부를 판단하는 블록도를 나타내고, (그림 7)은 글자가 삽입된 영상에서 글자가 삽입된 영역을 처리하는 과정을 나타낸다. 본 논문에서는 균질질감(Homogeneous Texture)에서 추출한 1~6번 특징값을 입력으로 하는 베이스 분류기(Bayesian Classifier)를 적용하여 글자가 삽입된 영상을 검출하였다. 이렇게 글자가 삽입된 영상으로 검출된 영상에 대해서는 글자가 삽입된 영역을 찾아서 제거하는 과정을 거친다. 글자가 삽입된 양이 적은 경우, 실제 검색률에 큰 영향을 미치지 못하므로 첫 번째 단계에서 글자가 있는 영상으로 검출되는 경우, 글자가 삽입된 영역이 넓게 분포함을 알 수 있다.

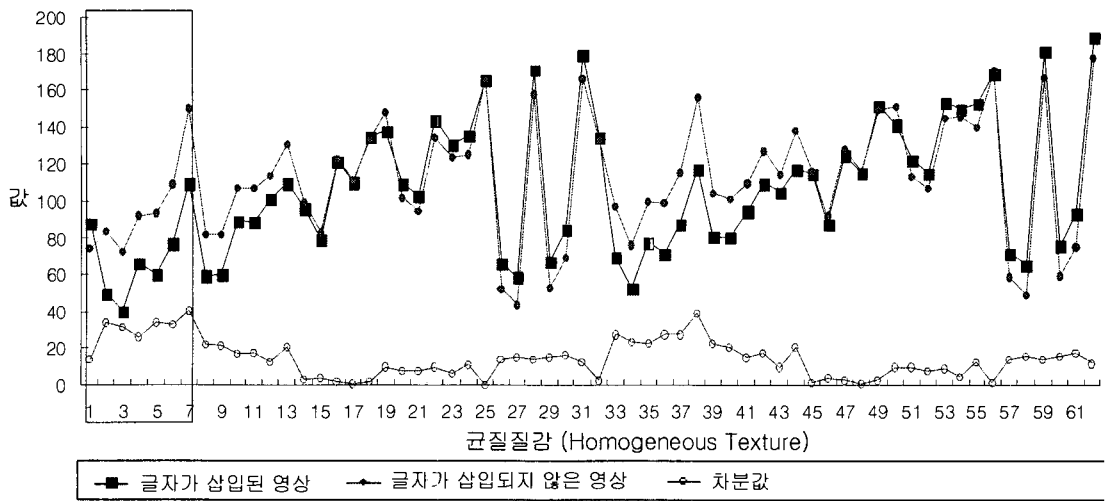


그림 5. 글자가 삽입된 영상과 글자가 삽입되지 않은 영상의 균질질감

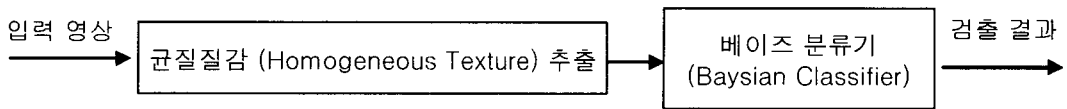


그림 6. 글자 삽입 여부 판단 블록도

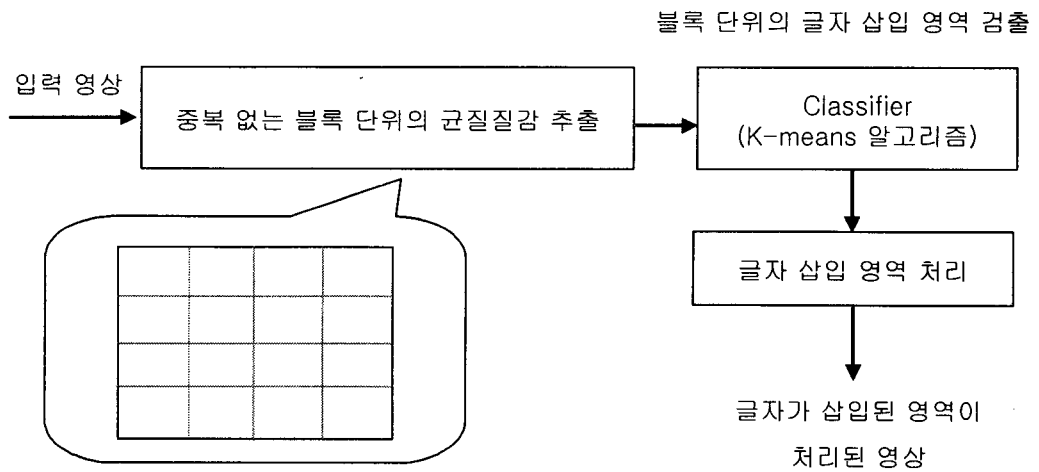


그림 7. 영상에 삽입된 글자 영역의 처리 블록도

글자가 삽입된 영상으로 판별된 영상에 대해서는 글자가 삽입된 영역과 삽입되지 않은 영역을 판별하고, 글자가 삽입된 영역에 대해서는 글자에 의한 질감특징을 감소하는 전처리를 수행하였다. 먼저 전체

영역을 가로 세로 각각 4개, 즉 16블록으로 나누고 각각의 블록에 대해 질감정보중 1~6의 특징값을 계산한다. 16개 블록은 이 특징값을 입력으로 글자가 삽입된 블록과 삽입되지 않은 블록으로 K-means 알

고리즘을 이용하여 클러스터링 된다. 글자가 삽입된 것으로 판별된 영역에 대해서는 11 tap의 평균화 필터를 통해 글자의 질감특성을 제거하는 과정이 글자 삽입 영역 처리 블록에서 수행된다.

3.2 프레임이 삽입된 영상의 편집효과 제거

프레임이 삽입된 영상의 전처리하는 프레임을 제거하여 프레임 삽입으로 인한 영향을 최대한 줄여주는 데 목적을 두고 있다. 프레임이 삽입된 영상은 프레임의 특성을 고려하여 다음과 같은 방법으로 전처리 과정을 수행한다.

프레임이 삽입된 영상은 외곽에 위치한 프레임과 프레임을 제외한 부분영상으로 구성되어 있다. 일반적으로, 외곽에 위치한 프레임과 중앙에 위치한 부분영상의 경계는 밝기값의 변화가 뚜렷하기 때문에 수직·수평 영사값(Projection Value)의 변화가 매우 크다. 따라서, 추출하고자 하는 부분영상과 프레임의 경계는 1차원의 영사값에 표준편차(Standard Deviation)를 적용하여 구할 수 있는데, 삽입된 프레임이 복잡한 형태일 수도 있다는 점을 고려하여 국부표준편차(Local Standard Deviation)를 적용한다. 또한 삽입된 프레임의 특성상 프레임이 외곽에 위치하기 때문에, 임계값과 계산된 국부표준편차의 비교 범위는 외

곽으로부터 영상의 수직·수평길이의 1/4 범위 내에서만 비교한다. 식 (1)은 수직·수평 영사값(Projection Value)을 나타낸다.

$$P_y(x) = \frac{1}{N} \sum_{y=0}^{N-1} f(x,y), \quad P_x(y) = \frac{1}{M} \sum_{x=0}^{M-1} f(x,y) \quad (1)$$

여기서  $f(x,y)$ 는  $(x,y)$ 위치에서의  $N \times M$ 크기를 가지는 흑백영상의 화소값이며,  $P_y(x)$ ,  $P_x(y)$ 는 각각, 수직, 수평으로의 영사값이다. 식(2)에서는 주위 4 픽셀을 이용한 국부표준편차(Local Standard Deviation) 값인  $S_x(y)$ ,  $S_y(x)$ 을 나타낸다.

$$S_x(y) = \sqrt{\frac{1}{5} \sum_{k=-2}^2 \{P_x(y+k)\}^2 - \left\{ \frac{1}{5} \sum_{k=-2}^2 P_x(y+k) \right\}^2},$$

$$S_y(x) = \sqrt{\frac{1}{5} \sum_{k=-2}^2 \{P_y(x+k)\}^2 - \left\{ \frac{1}{5} \sum_{k=-2}^2 P_y(x+k) \right\}^2} \quad (2)$$

식 (3)은 국부표준 편차  $S_x(y)$ ,  $S_y(x)$ 가 임계값보다 큰 프레임과 부분영상간의 경계선을 결정하는 수식으로,  $X_s$ ,  $X_e$ 는 프레임의 수직 경계선을,  $Y_s$ ,  $Y_e$ 는 프레임의 수평 경계선을 나타낸다. (그림 8)은 영사값의 국부표준편차와 추출된 경계를 보여준다.

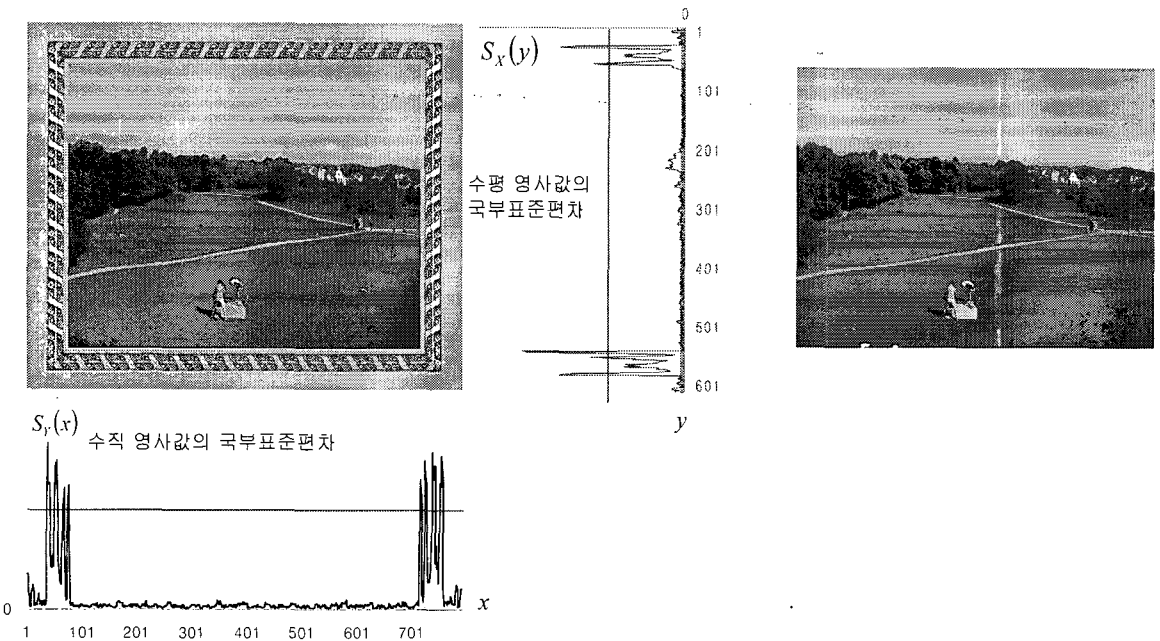


그림 8. 프레임이 삽입된 영상과 영사값의 국부표준편차

$$X_s \leq m \leq X_e, \quad Y_s \leq n \leq Y_e$$

$$X_s = \max_x \left\{ S_Y(x) \geq \text{Threshold}, 0 \leq x \leq \frac{1}{4}M \right\}$$

$$X_e = \min_x \left\{ S_Y(x) \geq \text{Threshold}, \frac{3}{4}M \leq x \leq M \right\}$$

$$Y_s = \max_y \left\{ S_X(y) \geq \text{Threshold}, 0 \leq y \leq \frac{1}{4}N \right\}$$

$$Y_e = \min_y \left\{ S_X(y) \geq \text{Threshold}, \frac{3}{4}N \leq y \leq N \right\} \quad (3)$$

프레임제거를 위한 전처리 과정은 다음과 같다.

**step 1:** 입력영상을 흑백영상으로 변환한다.

**step 2:** 흑백영상으로부터 수직·수평 영사값을 구한다.

**step 3:** 영사값의 국부표준편차를 구한다.

**step 4:** 추출될 부분영상의 수직, 수평의 프레임 경계 포인트의 위치를 구한다.

**step 5:** 입력영상으로부터 부분영상을 추출한다.

### 3.3 여러영상으로 구성된 영상의 편집효과 제거

여러영상으로 구성된 영상의 전처리는 전체영상을 단일 영상이라고 할 수 있는 각각의 부분영상들로 분할하여 각 부분영상들의 특징값이 중첩되는 것을 방지한다. 그리고 각각의 분할된 영상들 중에서 사용자 질의 영상과 유사한 영상이 하나 이상일 경우, 전체영상은 사용자 질의 영상과 내용이 유사한 영상으로 판단된다.

여러영상으로 구성된 영상은 부분영상들의 경계에서 직선이 검출되고, 검출된 직선에 의해서 닫히는 영역을 부분영상으로 판단할 수 있다. 그러나 상·하·좌·우로 1개의 직선이 검출되는 프레임과는 다르게 1개 이상의 직선이 검출될 수 있고, 부분영상들의 경계선의 길이가 짧다면 임계값을 이용한 판단여부가 어려워진다. 따라서, 비슷한 칼라나 질감을 가지는 1개의 프레임을 제거하는 것보다 여러영상을 부분영상으로 분리하기 위해 좀더 복잡한 알고리즘이 요구된다. 식 (4)는 수직·수평 직선,  $l_v(x,y)$ ,  $l_x(x,y)$ 을 구하기 위해 사용되는 마스크를 전개한 식을 나타내고 식 (5)는 검출된 수직·수평 직선  $e(m,y)$ ,  $e(x,n)$ 을 나타낸다.

$$l_v(x,y) = f(x,y) + f(x,y-1) + f(x,y+1) - f(x-1,y-1) - f(x-1,y) - f(x-1,y+1) \quad (4)$$

$$l_x(x,y) = f(x,y) + f(x-1,y) + f(x+1,y) - f(x-1,y-1) - f(x,y-1) - f(x+1,y-1)$$

$$e(m,y), \quad m \in \{v \mid S_Y(x) > \text{threshold}\} \quad (5)$$

$$e(x,n), \quad n \in \{h \mid S_X(y) > \text{threshold}\}$$

다음은 여러영상을 분해하는 과정을 보여준다.

**step 1:** 입력영상을 흑백영상으로 변환한다.

**step 2:** 흑백영상으로부터 수직·수평 직선을 구한다.

**step 3:** 수직 직선으로부터 수직 영사값과 국부표준편차를 구하고 수직 경계선을 파악한다. 그리고 같은 방법을 이용하여 수평 경계선을 파악한다.

**step 4:** (그림 9)의 (a)에서와 같이 수직라인에 의해서 나뉘어지는 4개의 수평 분할선의 수평 영사값과 수평영사값의 국부표준편차를 구하여 임계값보다 높은 수평 분할선을 선택한다. 그리고 같은 방법으로 (그림 9) (b)에서 수직 분할라인을 선택한다.

**step 5:** 분할된 수직·수평 라인들에 의해서 닫히는 영역의 좌표를 얻어 부분영상을 추출한다. (그림 10)은 4개의 분할라인과 3개의 닫힌 영상을 보여주고 있다.

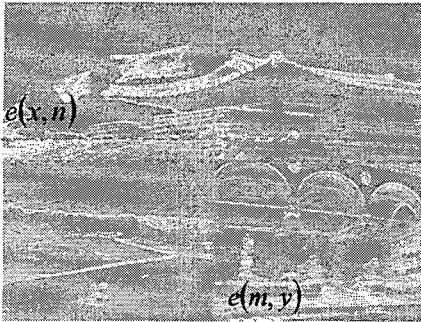
### 3.4 편집효과를 고려한 내용기반 영상 검색시스템

일반적으로 편집효과는 한 가지만 적용되기도 하지만, 하나 이상의 편집효과가 동시에 적용되는 경우도 발생한다. 따라서, 혼합된 편집을 모두 제거하기 위해서는 편집효과를 제거하는데 소요되는 시간과 편집효과를 제거하는 순서가 고려하여야 한다.

본 논문은 편집효과를 제거하는 시간을 최소화하기 위해 각각의 편집효과 제거 알고리즘을 편집효과의 검출(Detection)과 전처리(Preprocessing)로 분

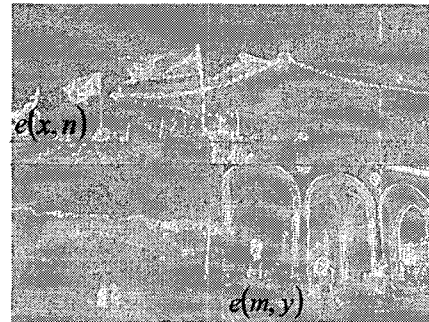


수직 라인 영상 :  $l_y(x, y)$



(a) 수직라인  $e(m, y)$ 에 의해서 분할되는 4개의 수평 분할선

수평 라인 영상 :  $l_x(x, y)$



(b) 수평라인  $e(x, n)$ 에 의해서 분할되는 3개의 수직 분할선

그림 9. 여러영상을 분해하기위한 수직, 수평분할선의 검출

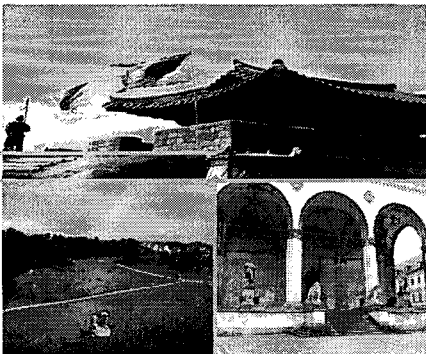


그림 10. 수직·수평 분할라인에 의한 4개의 분할라인과 3개의 닫힌영역

리하였다. 편집효과 검출은 입력영상에 대해 편집의 적용 유무를 확인하는 과정이고, 편집효과 전처리하는 확인된 편집 종류에 따라 프레임 제거, 부분영상으로의 분리, 글자 삽입 블록의 처리와 같은 편집효과 제거의 역할을 한다.

편집효과를 제거하는 순서는 적용된 편집효과 제거 알고리즘의 특성을 고려하여 결정되어야만 한다. 글자가 삽입된 영상은 전체영상을 블록 단위로 나누어 글자가 삽입된 블록을 평탄화한다. 그러나 평탄화된 블록은 영상 내 영상 (Picture-in-Picture)으로 판단되기 쉽기 때문에 삽입된 글자를 제거하는 알고리즘은 여러영상을 분해하는 알고리즘이 수행된 후에 적용되도록 하였다. 또한, 프레임이 삽입된 영상에서 프레임을 제거하고 원본영상만을 추출하는 알고리즘의 특성에 따라 프레임을 제거하는 알고리즘은 여

러영상을 분해하는 알고리즘보다 먼저 수행되도록 결정하였다. 따라서, 편집효과를 제거하는 시간과 편집효과를 제거하는 순서를 고려한다면 (그림 11)와 같이 편집효과 제거 알고리즘을 구성할 수 있고, 이를 적용한 내용기반 영상 검색시스템은 (그림 12)과 같이 설계될 수 있다.

#### 4. 실험 및 분석

제안한 전처리가 적용되었을 때 내용기반 검색시스템의 성능향상을 확인하기 위해 유해영상 필터링 시스템에 제안한 방법을 적용하였다. 이 때 사용된 내용기반 필터링 시스템은 (표 1)의 결과를 얻는 데 사용한 MPEG-7의 에지, 칼라히스토그램, 균질질감을 입력으로 받는 베이시안 분류기이다. 사용된 유해영상은 편집되지 않은 유해영상과 편집된 유해영상으로 구분되고, 편집된 유해영상은 글자가 삽입된 영상, 프레임이 삽입된 영상, 여러영상으로 구성된 영상으로 다시 세분된다.

(표 3)은 전처리에서의 편집효과 검출성능을 나타낸다. 실험결과에서 볼 수 있듯이 세 가지 편집효과 검출률이 84~89%를 나타낸다. 실험에서 검출되지 않은 편집효과들은 삽입된 글자의 양이 적거나, 삽입된 프레임의 형태가 사각의 테두리가 아닌 복잡한 형태인 경우였다. 여러영상으로 구성된 영상 가운데 잘 검출되지 않는 경우는 부분영상들의 경계가 불명확한 특징을 가지고 있었다.

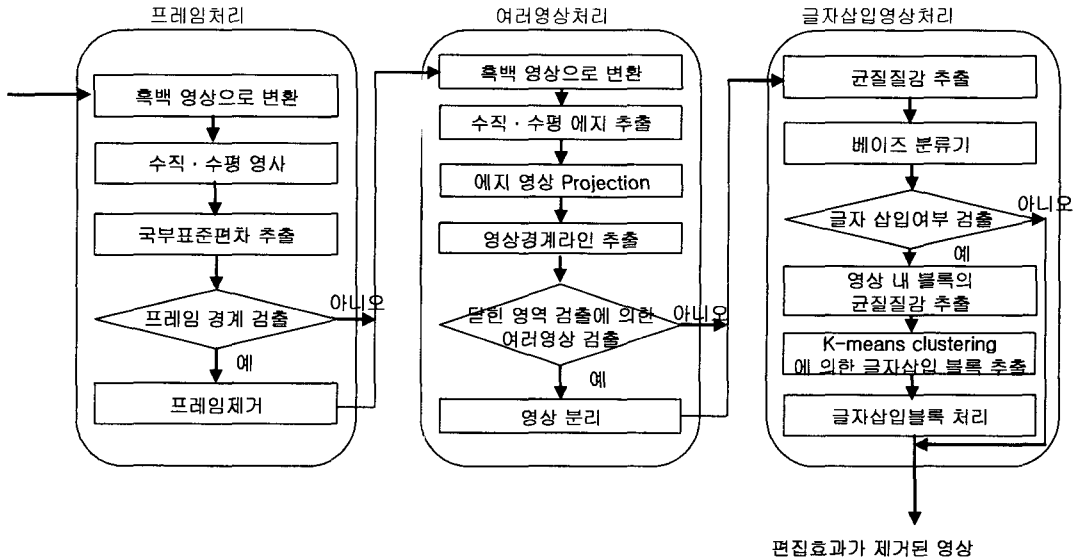


그림 11. 편집효과 제거 알고리즘의 특성을 고려한 편집효과 제거 과정

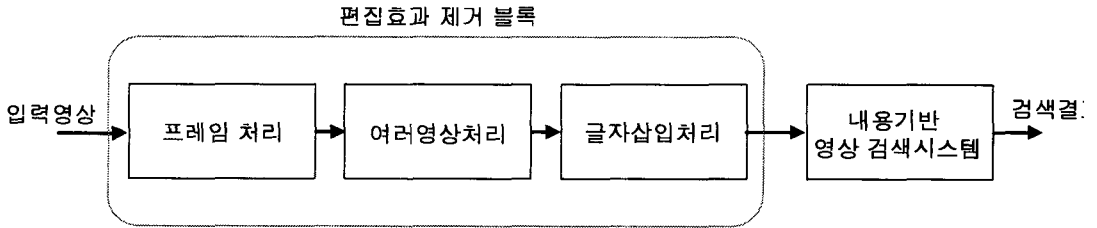


그림 12. 향상된 내용기반 영상 검색시스템

표 3. 전처리기에서 각 편집효과에 대한 검출률

편집효과	리콜률	정확도(%)
글자 삽입	0.84	95.4
프레임 삽입	0.85	98.8
여러영상으로 구성	0.89	95.7

표 4. 제안한 검색시스템의 성능 향상 결과

편집된 영상	리콜률	
	전처리 전	전처리 후
글자가 삽입된 영상	0.13	0.37
프레임이 삽입된 영상	0.33	0.55
여러영상으로 구성된 영상	0.52	0.72

(표 4)는 제안한 전처리 알고리즘이 적용된 유해 영상 필터링시스템의 성능 향상을 보여주고 있다. 필터링 성능은 전처리가 적용되지 않았을 때 보다 약

24% 정도 향상되는 것을 확인할 수 있다. 이는 시각적으로 편집효과를 줄여주는 전처리 알고리즘이 내용기반 검색시스템의 성능향상을 보장하는 것을 의

미한다. 또한 수작업을 통한 시각적 전처리 결과를 보여주는 (표 2)와 비교하였을 때, 글자가 삽입된 영상에 대한 전처리 알고리즘이 좀 더 보완되어야 함을 알 수 있다.

## 5. 결 론

본 논문에서는 내용기반 검색시스템의 성능을 저하시키는 영상 편집효과를 분석하였으며, 이러한 영상편집효과들이 영상의 특징정보인 색상, 질감, 형태 등을 왜곡시켜 내용기반 검색시스템의 성능을 저하시키는 것을 확인하였다. 이러한 문제를 해결하기 위해, 영상의 편집효과를 검출하고 제거하는 전처리 알고리즘과 제안된 전처리 알고리즘을 적용한 내용기반 영상 검색시스템을 제안하였으며, 내용기반 필터링 시스템에 적용하여 제안한 시스템의 유용성을 확인하였다. 그러나 복잡한 형태의 프레임이 삽입된 영상에 대한 검출률을 개선하고 그림 속 그림(Picture-in-Picture) 형태로 삽입된 영상과 삽입된 글자검출에 대해 좀 더 정밀한 전처리 방법에 대한 연구가 필요하다. 또한, 대용량 영상 검색시스템에 적용하기 위해서는 연산량이 작은 전처리 알고리즘에 대한 연구가 요구된다.

## 참 고 문 헌

- [ 1 ] ISO/IEC JTC1/SC29/WG11 N4980, "MPEG-7 Overview," July 2002.
- [ 2 ] ISO/IEC 15938-3, "Information Technology Multimedia Content Description Interface Part 3: Visual," July 2002.
- [ 3 ] ISO/IEC JTC1/SC29/WG11MPEG02/N4582, "MPEG-7 Visual part of eXperimentation Model Version 13.0," March 2002.
- [ 4 ] Manjunath, Philippe Salembier, and Thomas Sikora, "Introduction to MPEG-7 Multimedia Content Description Interface, B.S.," *John Wiley & Sons, LTD*, 2002.
- [ 5 ] Kyoji Hirata and Toshikazu Kato, "Rough Sketch-Based Image Information Retrieval," *NEC Res. & Develop.*, Vol. 34, No. 2, 1993.
- [ 6 ] 염성주, 김우생, "형태와 컬러성분을 이용한 내용 기반의 이미지 검색 데이터베이스 시스템," *한국정보처리학회 논문지*, 제3권 4호, pp. 733-744, July 1997.
- [ 7 ] M. Flickner, et al., "Query by Image and Video Content: The QBIC System," *IEEE Computer*, Vol. 28, No. 9, pp 23-31, Sept. 1995.
- [ 8 ] J. R. Smith and S.-F. Chang, "VisualSEEK: a fully automaed content-based image query system," *In Proc. ACM Intern. Conf. Multimedia*, Boston MA, pp. 87-98, November 1996, <http://www.ctr.columbia.edu/VisualSEEK>
- [ 9 ] 이동호, 송용준, 김형주, "SCARLET: 웨이블릿 변환을 이용한 내용기반 이미지 검색 시스템의 설계 및 구현," *한국정보과학회 논문지(C)*, 제3권 4호, pp. 353-364, August 1997.
- [10] A. Antonacopououlos, D. Karatzas, and J. Ortiz Lopez, "Accessing Textual Information Embedded in Internet Images," *Proceedings of SPIE, Internet Imaging II*, SanJose, USA, January 2001, Vol. 4311, pp. 198-205.
- [11] G. Murch, "Color /Displays and Color Science," *in Color and the Computer*, J. H. Durrett, Ed. Orlando, Florida: Academic Press INC., 1987, pp. 1-25.
- [12] A.K. Jain and B. Yu, "Automatic Text Location in Images and Video Frames," *Pattern Recognition*, Vol. 31, No. 12, 1998, pp. 2055-2076.
- [13] J. Gllavata, R. Ewerth, and B. freisleben, "A Robust Algorithm for Text Detection in Images," *Proceedings of the 3<sup>rd</sup> International Symposium on Image and Signal Processing Analysis*, pp. 611-616, May 2003.
- [14] L. Agnihotri and N. Dimitrova. "Text Detection for Video Analysis," *In Proc. Of the International Conference on Multimedia Computation and Systems*, Florence, Italy, pp. 109-113, 1999.
- [15] Kwang In Kim, Keechul Jung, and Jin Hyung Kim, "Texture-Based Approach for Text Detection in Images Using Supprot Vector Machines and Continuously Adaptive Mean Shift Algorithm," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 12, pp. 1631-1638, Dec. 2003.



**강 석 준**

2003년 경북대학교 전자전기공학부 공학사  
 2005년 한국정보통신대학교 공학부 공학석사  
 관심분야 : Digital Watermarking, Retrieval, Video Coding



**배 태 면**

1996년 2월 경북대학교 전자공학과 공학사  
 1998년 2월 경북대학교 전자공학과 공학석사  
 2000년 8월 경북대학교 박사과정 수료  
 2000년 9월~2001년 3월 Togabi

기술 연구원

2001년 3월~현재 서울대학교 컴퓨터기술연구소 연구원

2002년 9월~현재 한국정보통신대학교 박사과정

관심분야 : semantic video analysis, Scalable Video Coding, Video Adaptation, Watermarking



**김 기 현**

1989년 2월 경북대학교 전자공학과 공학사  
 1991년 2월 경북대학교 컴퓨터공학과 공학석사  
 1991년 2월~1998년 12월 한국전자통신연구원  
 1999년 1월~2002년 4월 하이퍼

정보 주식회사

2002년 5월~현재 한국전자통신연구원

2000년 3월~현재 한국정보통신대학교 박사과정

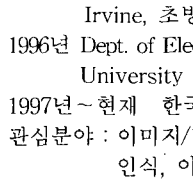
관심분야 : Video Codec Design, Scalable Video Coding, Watermarking, Biometric Image Analysis



**노 용 만**

1985년 2월 연세대학교 전자공학과 공학사  
 1987년 2월 KAIST 전기공학과 공학석사  
 1992년 2월 KAIST 전기공학과 공학박사  
 1992년~1995년 Dept. of

Radiological Science, University of California,



**한 승 완**

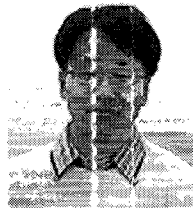
1994년 전남대학교 전산학과 이학사

1996년 전남대학교 전산통계학과 이학석사

2001년 전남대학교 전산통계학과 이학박사

2001년~현재 한국전자통신연구원 정보보호연구단 정보보호원천연구팀 선임 연구원

관심분야 : 네트워크 보안, 알고리즘, 계산 이론 등



**정 치 윤**

2002년 포항공과대학교 전자전기공학과 공학사.

2004년 포항공과대학교 전자전기공학과 공학석사.

2004년~현재 한국전자통신연구원 정보보호연구단 능동 보안기술연구팀 연구원.

관심분야 : 컴퓨터 비전, 영상 처리, 네트워크보안 등



**남 택 용**

1987년 충남대학교 계산통계학과 이학사

1990년 충남대학교 계산통계학과 이학석사

2005년 한국외국어대학교 전자정보공학과 공학박사

1987년~현재 한국전자통신연구원 정보보호연구단 보안게이트웨이연구팀 팀장(책임연구원). 관심분야는 정보보호, 인터넷, 이미지 마이닝 등