

# 비디오에서 불투명 및 반투명 TV 로고 인식을 위한 로고 전이 검출 방법

## (A Logo Transition Detection Method for Opaque and Semi-Transparent TV Logo Recognition in Video)

노 명 철 <sup>†</sup>                      강 승 연 <sup>\*\*</sup>                      이 성 환 <sup>\*\*\*</sup>  
 (Myung-Cheol Roh)      (Seung-Yeon Kang)      (Seong-Whan Lee)

**요 약** UCC(User Created Contents)의 급격한 증가에 따라 저작권 문제도 크게 대두되고 있다. 자동 로고 인식은 이러한 저작권 문제를 해결하기 위한 효율적인 방법이다. 로고는 다양한 특징을 가지고 있고, 이러한 특징들은 로고 검출과 인식을 어렵게 한다. 특히, 비디오 내에 빈번한 로고 전이가 일어날 경우, 정확한 로고 인식과 로고 기반 분할이 어렵다. 따라서 본 논문에서는 디지털 비디오에서 로고 인식을 위한 정확한 전이 검출 방법과 다양한 로고 타입 인식 방법을 제안한다. 제안한 로고 검출과 로고에 따른 비디오 분할을 이용하여 다양한 비디오에 대한 좋은 실험 결과를 얻을 수 있었다.

키워드 : 로고 전이 검출, 로고 검출 및 인식, 부-비디오 병합, 로고 기반의 비디오 분할

**Abstract** The amount of UCCs (User Created Contents) has been increasing rapidly and is associated with a serious copyright problem. Automatic logo detection in videos is an efficient means of overcoming the copyright problem. However, logos have varying characteristics, which make logo detection and recognition very difficult. Especially, there are frequent logo transitions in a video, comprising several video contents. This disrupts accurate video segmentation based on logos. Therefore, this paper proposes an accurate logo transition detection method for recognizing logos in digital video contents. The proposed method accurately segments a video according to logo and efficiently recognizes various types of logos. The experimental results demonstrate the effectiveness of the proposed method for logo detection and video segmentation according to logo.

**Key words** : Logo transition detection, logo detection and recognition, sub-video merging, video segmentation based on logos

### 1. 서 론

계속되는 인터넷의 발전과 함께 늘어나는 UCC(User Created Contents)의 비디오 콘텐츠 유통은 심각한 저작권 문제를 야기하고 있다. 비디오의 원저작자는 본인 소유의 비디오에 대한 저작권을 행사하고자 하지만, 사용자들이 불법적이거나 합법적으로 모은 디지털 비디오들을 저작권자의 허락 없이 임의로 조작하고 유통시키는 것이 저작권 문제를 심각하게 만들고 있다. 비디오의 로고는 원저작자의 저작권을 표시하는 것으로서, 비디오에서의 자동 로고 검출은 불법적인 저작권문제를 해결하기 위한 효율적인 수단의 하나이다. 또한 자동 로고 검출 기술은 비디오 서술과 색인[1,2], 로고 제거[3-6], 광고와 TV 프로그램 구별[7-9], 스포츠 경기의 리플레이(Replay) 검출[10] 등 다양한 응용분야가 있다.

로고 검출과 인식에는 여러 가지 해결해야 할 문제들

· 이 연구에 참여한 연구자는 '2단계 BK21 사업'의 지원비를 받았다

<sup>†</sup> 학생회원 : 고려대학교 컴퓨터-통신공학부  
 mroh@image.korea.ac.kr

<sup>\*\*</sup> 비 회 원 : 고려대학교 컴퓨터-통신공학부  
 sykang@image.korea.ac.kr

<sup>\*\*\*</sup> 종신회원 : 고려대학교 컴퓨터-통신공학부 교수  
 swlee@image.korea.ac.kr  
 (Corresponding author인)

논문접수 : 2008년 10월 6일

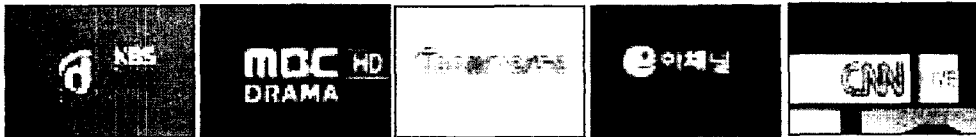
심사완료 : 2008년 11월 6일

Copyright©2008 한국정보과학회 : 개인 목적이거나 교육 목적인 경우, 이 서적물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

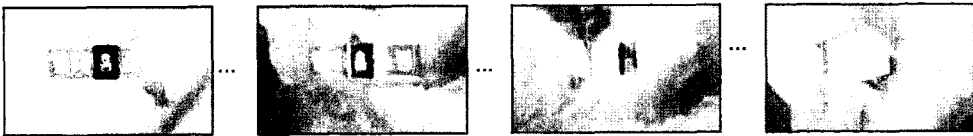
정보과학회논문지: 소프트웨어 및 응용 제35권 제12호(2008.12)



(a) 반투명 로고의 예



(b) 불투명 로고의 예



(c) 애니메이션 로고의 예  
그림 1 다양한 로고 타입

이 있다. 첫 번째 문제로는 TV 로고는 색, 모양 등의 다양한 특징을 가진다는 것이다. 이 로고는 반투명 로고, 불투명 로고, 애니메이션 로고로 3가지 종류로 나눌 수 있다. 그림 1은 이러한 로고들의 예를 보여준다. 로고의 배경이 로고를 통하여 보이면 반투명 로고라 하고, 배경이 보이지 않으면 불투명 로고라 한다. 애니메이션 로고는 로고 자체가 회전하거나 움직이는 로고이다. 불투명 로고는 배경의 변화에 독립적이고, 반투명 로고는 배경의 변화에 영향을 받는다. 이러한 성질로 인하여 비디오 프레임에서 반투명 로고의 추출이 어렵다. 그러므로 시간 영역에서의 로고 검출 방법들이 일반적으로 사용된다.

로고 검출에서 두 번째 문제는 비디오에서 발생하는 로고 전이(Transition)이다. 로고 전이에는 로고의 나타남(비디오의 중간에 없던 로고가 나타나기 시작함), 로고의 사라짐(비디오의 중간에서 로고가 없어짐), 로고의 바뀜(비디오 중간에 로고 A에서 로고 B로 바뀜)이 있을 수 있다. 그림 2는 이러한 예를 보여주고 있다. 이러한 로고 전이는 UCC와 같이 사용자 임의로 변경하는 비디오에서 자주 나타난다. 그러므로 로고의 검출을 위해서는 로고 전이 검출이 필수적이다.

기존의 로고 검출 방법은 그 응용에 따라서 단일 비디오 프레임 기반 검출과 연속된 시간 공간상의 비디오 프레임 기반 검출로 나누어진다.

Seebeer 등은[11] TV에서 반투명 로고 추출을 위하여 내부와 외부 에지를 사용한 단일 비디오 프레임의

형판 정합(Template Matching) 방법을 제안하였다. 이 방법은 고해상도에서만 효율적으로 동작한다는 단점이 있다. Duffner 등은[12] 이 방법을 개선하여 Convolved Neural Network 기반의 다중 해상도 방법을 제안하였다.

일반적으로는 연속된 비디오 프레임을 사용하는 시간 공간 정보를 사용한 로고 검출 방법들이 많이 사용된다. Yan 등은[3] 비디오 영상 간의 차이로부터 로고를 검출하여 로고를 제거하는 방법을 제안하였다. 이 방법은 컬러 정보를 포함하는 로고 추출에 단점을 보인다. Bai 등은[13] 움직임 분석 기반의 로고 검출 방법을 제안하였다. 동적 프로그래밍과 비교사 클러스터링을 이용하여 로고 형판이 학습되고, 광류(Optical Flow) 특징을 이용하여 로고 변화를 검출하여 스포츠 경기 비디오에서 리플레이를 검출하였다. Cozar 등은[1] 비디오 색인을 위한 구성을 제안하였다. 최소 휘도 분산 영역(Minimum Luminance Variance Region: MVLV)을 추출하여 시공간적인 분할을 수행하여 로고를 검출하고 인식한다. 이 방법은 반투명 로고 검출에 효율적이나, 로고 전이가 있는 비디오에는 정확한 로고를 인식할 수 없다는 단점이 있다. 본 논문에서는 장면 검출을 사용하여 로고 전이를 검출할 수 있는 방법과 불투명 및 반투명 로고를 분류하여 검출 및 인식하는 방법을 제안한다.

## 2. 로고 전이 검출 기반의 로고 검출 및 인식 방법

제안하는 로고 전이 검출 기반의 로고 검출 및 인식 방법은 크게 5가지의 구성 성분으로 이루어진다. 그림

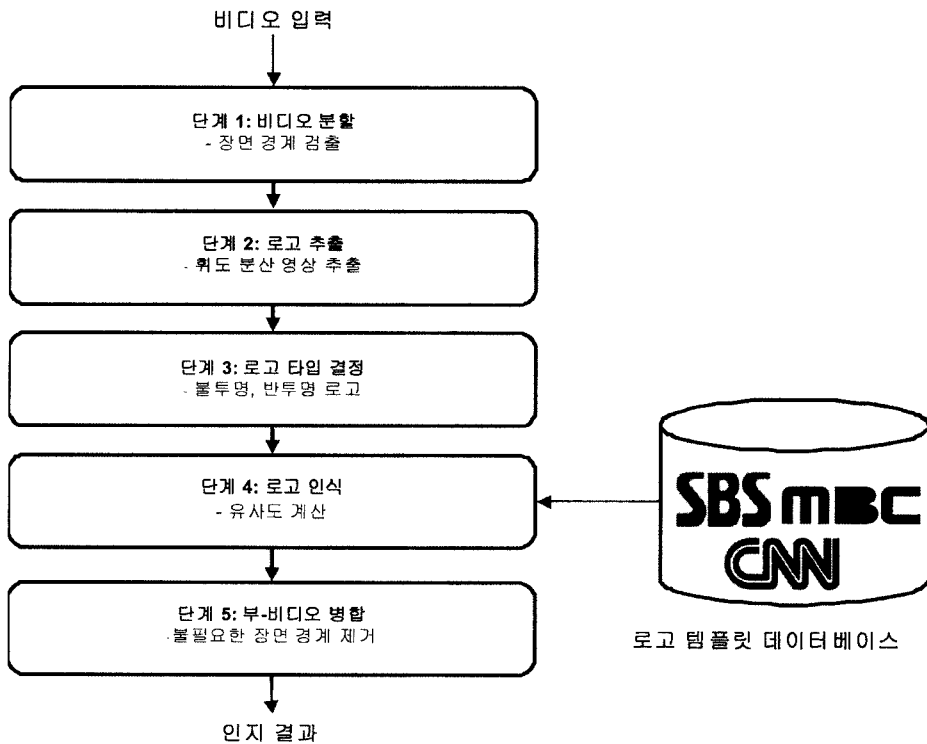


그림 3 제안한 로고인식 시스템의 흐름도

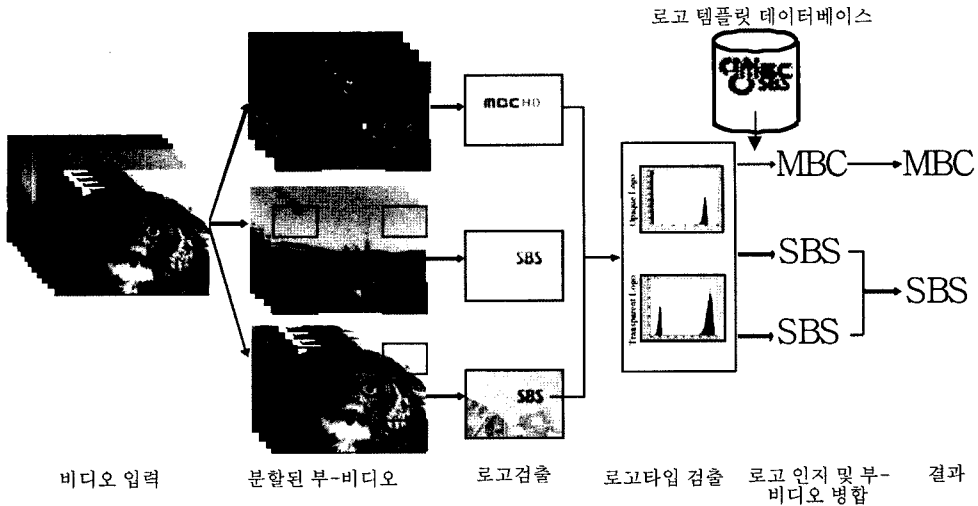


그림 4 제안하는 로고 인식 방법의 과정

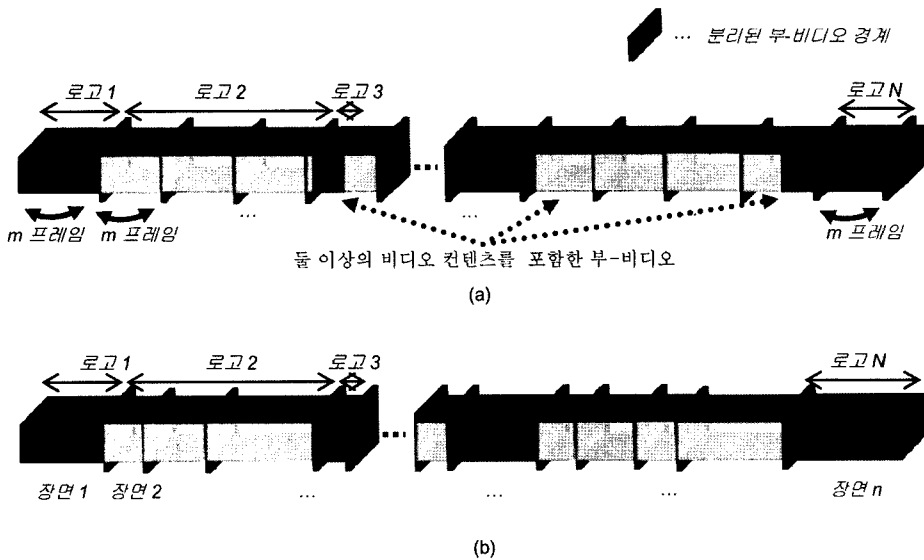


그림 5 (a) 고정된 프레임 길이를 사용한 비디오 분할 방법, (b) 제안하는 장면 경계 검출 기반 비디오 분할 방법

3, 4는 본 논문에서 제안하는 로고 전이 검출 기반의 로고 검출 및 인식 방법을 보여준다. 입력 비디오는 장면 검출을 이용하여 여러 개의 부-비디오(Sub-Video)로 나누어진다. 각 부-비디오의 연속된 프레임에서 휘도 분산을 계산하여 후보 로고 형판을 추출한다.

추출된 후보 로고가 불투명 인지, 반투명 인지를 판단하고 적합한 방법으로 인식한다. 불필요한 부-비디오의 경계를 없애 정확한 로고 전이의 경계를 찾기 위하여 인식된 로고를 기반으로 하여 부-비디오 간의 병합이 이루어진다. 그림 4는 두 개의 다른 로고를 가지는 비

오가 합쳐진 하나의 비디오로부터 제안된 방법을 이용하여 로고 전이를 검출하고 인식하는 간략한 과정을 보여준다.추출된 후보 로고가 불투명 인지, 반투명 인지를 판단하고 적합한 방법으로 인식한다. 불필요한 부-비디오의 경계를 없애 정확한 로고 전이의 경계를 찾기 위하여 인식된 로고를 기반으로 하여 부-비디오 간의 병합이 이루어진다. 그림 5는 두 개의 다른 로고를 가지는 비디오가 합쳐진 하나의 비디오로부터 제안된 방법을 이용하여 로고 전이를 검출하고 인식하는 간략한 과정을 보여준다.

**2.1 비디오 분할(Video Segmentation)**

앞에서 언급한 것과 같이 임의의 사용자에게 의하여 편 집된 비디오에서 로고를 찾기 위해서는 로고 전이 검출 이 필요하다. 기존의 방법에서는 주어진 길이의 비디오 단위로 분할하여 처리하므로 그림 5(a)에서 보는 것과 같이 잘못된 분할이 발생하고 그에 따라 잘못된 인식 결과를 얻게 된다. 그림 5(a)는 주어진 길이의 단위로 나누어서 처리할 경우 하나의 부-비디오 내에 서로 다른 로고들이 섞여서 분할되는 결과를 보여준다. 이 문제를 해결하기 위하여 로고 추적 기술도 제안되었지만, 비슷한 로고가 연속되는 경우에는 잘못된 결과를 얻게 된다. 그러므로 본 논문은 장면 검출 기법을 기반으로 한 비디오 분할 방법을 제안한다. 하나의 비디오는 장면 검출 기법으로 여러 개의 부-비디오로 나누어진다. 여기서 로고 전이가 발생하면, 비디오 장면이 전환된다는 가정을 한다. 그림 5(b)에서 보는 바와 같이, 하나의 로고를 가지는 비디오가 여러 개의 부-비디오로 나누어 질 수 있고, 여러 개의 로고를 가지는 부-비디오는 나타나지 않는다. 또한, 여러 개로 나누어진 장면 경계는 로고 전이를 포함하게 된다.

장면 경계 전환 검출을 위한 컷에는 급작스러운 컷(Abrupt Cuts)과 점진적인 컷(Gradual Cuts)의 두 가지가 있다. 로고 전이 시에는 비디오 내용, 배경 등의 변화로 인하여 급작스러운 컷을 수반한다. 그러므로 본 논문에서는 장면 경계 검출을 위해 널리 사용되는 연속되는 프레임간의 차이를 기반으로 한 급작스러운 컷 검출 방법을 이용한다.

Fang 등과 Boussaid 등은[4,14] 장면 경계 검출을 위한 일반적인 특징들을 설명하였다. 우리는 그 중에서 색상 히스토그램의 차이와 움직임 정보의 차이의 두 특징을 사용하였다.

첫 번째 특징은 연속된 프레임 간의 색상 히스토그램 차이(Color Histogram Difference: CHD)를 이용하는 것이다. CHD는 아래와 같이 정의된다.

$$CHD_i = 1 - \frac{1}{3n} \left[ \sum_{j=1}^{256} \min(R_i^j, R_{i-1}^j) + \sum_{j=1}^{256} \min(G_i^j, G_{i-1}^j) + \sum_{j=1}^{256} \min(B_i^j, B_{i-1}^j) \right] \quad (1)$$

위의 식에서  $n$ 은 프레임의 픽셀 개수를 나타내고,  $R_i^j, G_i^j, B_i^j$ 는 각각  $i$ 번째 프레임( $I_i$ )에서 Red, Green, Blue색 히스토그램의  $j$ 번째 빈(Bin)의 값들을 나타낸다. CHD<sub>*i*</sub>의 값은 0에서 1의 값을 가진다. 1에 가까울수록 두 프레임간의 차이가 커짐을 의미한다.

두 번째 특징은 움직임 보상 특징으로, 두 프레임간의 시각적인 불연속성을 측정할 수 있도록 한다. 먼저, 두 프레임은  $N \times N$ 개 픽셀의 블록으로 나누어진다. 본 논

문에서는  $N=16$ 을 사용하였다. 두 프레임간의 부동성은 다음과 같이 절대 차이의 합으로 측정된다.

$$sd(u, v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} |I_i(x, y) - I_i(x+u, y+v)| \quad (2)$$

위의 식에서  $I_i(x, y)$ 는  $i$ 번째 프레임의  $(x, y)$ 의 픽셀 값을 나타내고,  $N$ 는 블록의 크기를 나타낸다. 최소의  $sd$  값을 가지는 블록이 가장 잘 정합되는 블록으로 선택된다. 움직임 보상 특징은 아래와 같이 정의된다.

$$M_{N_i} = \frac{1}{N} \sum_{n=1}^{N_i} (|\bar{Y}_n - \bar{Y}_{n-1}| + |\bar{U}_n - \bar{U}_{n-1}| + |\bar{V}_n - \bar{V}_{n-1}|) \quad (3)$$

위의 식에서  $N_i$ 는 블록의 개수를,  $\bar{Y}, \bar{U}, \bar{V}$ 는 블록에서  $Y, U, V$  성분의 평균을,  $\bar{Y}', \bar{U}', \bar{V}'$ 는 가장 잘 정합되는 블록의 성분을 의미한다.

급작스러운 장면 컷은 위 두 특징들 중에서 피크(Peak)가 검출될 때 생성된다. 피크를 검출하기 위하여서  $i-2, i-1, i, i+1$  4개의 프레임을 사용하고, 피크는 아래의 조건이  $\Delta_1^k > 0$ 이고  $\Delta_2^k > 0$  을 만족할 때 검출된다.

$$\Delta_1^k = F_i^k - F_{i-1}^k, \quad \Delta_2^k = F_{i+1}^k - F_i^k \quad (4)$$

위의 식에서  $F_i^k$ , ( $k \in \{1, 2\}$ )는 식 (1)과 (3)에서 추출한 특징 값을 나타낸다.

결과적으로, 하나의 입력된 비디오는 여러 개의 부-비디오로 나누어지고, 각 부-비디오는 최대한 한 개의 로고를 가지게 된다.

**2.2 로고 추출**

비디오 분할 후에 얻어진 각 부-비디오의 연속되는 시-공간의 프레임들에서 로고 형판을 추출한다. 로고는 크게 애니메이션 로고와 정적 로고로 나누어지고, 정적 로고는 반투명과 불투명 로고로 나누어진다. 본 논문에서는 정적 로고 추출을 대상으로 한다. 애니메이션 로고는 계속되는 로고의 변화가 있기 때문에 다른 방법이 필요하다. 정적 로고는 여러 가지 특징을 가진다. 첫 번째, 같은 소유자의 TV 로고는 같은 모양과 크기를 가진다. 두 번째, 시간 영역에서 TV 로고의 분산은 배경 로고의 분산 보다 작다. 세 번째, 로고는 특정 위치에서 검출된다. 로고는 일반적으로 비디오의 좌상단, 좌하단, 우상단, 우하단에서 나타난다. 네 번째, 로고는 적은 수의 색이나 단일 색으로 이루어진다. 마지막으로, 로고는 다른 비디오와 구별되도록 배경과의 대조를 이룬다. 이러한 특징들을 기반으로 하여 로고 추출 방법을 제안한다.

2.2.1 휘도 분산 영상(Luminance Variance Image) 계산  
부-비디오( $\{S_1, S_2, \dots, S_n\}$ )로 이루어진 비디오를  $V = \{S_1, S_2, \dots, S_n\}$ 라고 하자. 각 부-비디오( $S_i, 1 \leq i \leq n$ )

는 연속되는 프레임( $\{f_1, f_2, \dots, f_k\}$ ,  $k_i$ 는 부-비디오의 프레임 개수)으로 이루어진다. 부-비디오는 이전에 언급한 장면 경계 추출 기법에 의해서 적출되고, 로고는 4개의 ROI(Regions of Interest)에서 추출된다. 로고 형편을 추출하기 위해서 부-비디오에서 최대 휘도 분산을 계산하는 LVI(Luminance Variance Image)의 방법을 사용한다[1]. 로고의 휘도 분산 값은 다른 영역의 값보다 작기 때문에, 최소 LVI 값을 가지는 픽셀들이 로고 후보 영역 픽셀이 된다.  $S_i(k)$ 를 부-비디오의  $k$ 번째 픽셀 값들의 집합이라고 하자.  $i$ 번째 부-비디오의 LVI는 아래의 수식으로 계산된다.

$$LVI_i(k) = \max(S_i(k)) - \min(S_i(k)), \quad 1 \leq k \leq p \quad (5)$$

위의 식에서  $k$ 는 픽셀의 인덱스이고  $p$ 는 하나의 프레임의 픽셀 개수이다.

부-비디오에서 로고 후보( $CL_i$ )는 다음과 같이 추출된다.

$$CL_i(k) = \begin{cases} LVI_i(k) & \text{if } LVI_i(k) \leq \tau \\ 0 & \text{if } LVI_i(k) > \tau \end{cases}, \quad 1 \leq i \leq p \quad (6)$$

위의 식에서  $\tau$ 는 LVI의 픽셀들의 평균값에 의해 결정되는 임계값이다. 평균 픽셀값이 높으면, 높은 임계값을 가진다.

### 2.3 로고 타입 결정

반투명과 불투명 로고는 서로 다른 특징을 가지고 있고, 인식을 위하여서는 다른 특징 추출 방법을 적용해야

하기 때문에, 로고 인식 전에, 후보 로고들의 타입을 결정해야한다.

반투명 로고의 휘도 분산 값인 불투명 로고의 값보다 크다. 불투명 로고의 휘도 분산 값은 0에 가까운 반면에 반투명 로고의 휘도 분산 값은 상대적으로 1의 값으로 치우친 결과를 보여준다. 이러한 차이를 이용하여 후보 로고가 반투명 로고인지 불투명 로고인지를 결정할 수 있다.

베이지 규칙을 이용하여 추출된 특징에 대한 로고 타입의 확률은 다음과 같이 계산된다.

$$P(LT|F) = kP(F|LT)P(LT) \quad (7)$$

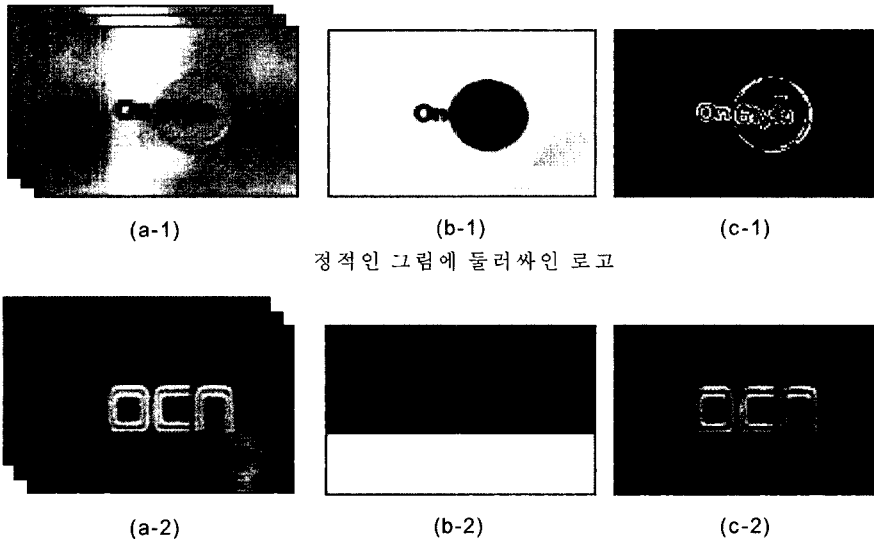
위의 식에서  $LT, F, k$ 는 각각 로고 타입, 특징, 그리고 정규 인자를 나타낸다. 후보가 반투명 로고일 확률과 불투명 로고일 확률은 아래와 같이 가정한다.

$$P(LT = \text{"Semi-transparent"}) = P(LT = \text{"Opaque"}) \quad (8)$$

$P(F|LT)$ 의 확률 분포는 학습 데이터를 이용하여 가우시안 모델로 모델링한다.

### 2.4 부-비디오에서 특징 추출과 로고 추출

로고 추출을 위한 특징 추출에 있어서, LVI의 에지 영상을 사용하는 것이 적합하지 않는 부-비디오가 있을 수 있다. 그림 6에서 보여주는 것과 같이 몇 개의 비디오에서는 로고 안의 에지나 로고 자체의 에지가 사라지는 경우가 생긴다. 그림 6의 첫 번째 열은 입력 영상을 보여주고(a-1, a-2) 두 번째와 세 번째 그림은 Cozar의



정적인 그림에 둘러싸인 로고

정적인 배경위에 나타나는 로고

그림 6 정적인 그림에 둘러싸인 로고 문자열을 가지는 로고의 예(윗줄)와 정적인 배경위에 나타나는 로고의 예(아랫줄): (a-1), (a-2) 입력 비디오 프레임, (b-1), (b-2) Cozar 등의 방법[1]으로 추출된 특징, (c-1), (c-2) 제안된 방법으로 추출된 특징

방법과 제안한 방법에 의한 특징 영상을 보여준다. 그림 6(a-1)은 원안에 텍스트가 싸여져 있는 로고의 예를 보여주고 있고, 그림 6(b-1)은 정적인 배경에 의해서 로고가 싸여진 예를 보여주고 있다. 이러한 비디오의 경우 이전의 방법들로는 특징 추출에 실패하고 결과적으로 인식이 실패하게 된다. 그러므로 우리는 로고 타입에 기반한 특징 추출방법을 제안한다.

만약 로고 형판이 불투명 로고라면 LVI의 에지와 부-비디오에서 임의로 선택된 프레임의 조합으로 특징을 추출한다. 불투명 로고의 에지는 모든 프레임에서 깨끗하게 추출될 수 있기 때문에 임의의 프레임을 참조 영상으로 사용할 수 있다. 불투명 로고 형판의 특징은 다음과 같이 계산된다.

$$L(i) = \begin{cases} \max(E_R(i), E_{CL}(i)) & \text{if } CL(i) = 0 \\ E_{CL}(i) & \text{otherwise} \end{cases} \quad (9)$$

위의 식에서  $E_R(i)$ 와  $E_{CL}(i)$ 는 참조 영상과 로고 후보 형판의 에지 영상들에서  $i$ 번째 픽셀값을 각각 나타낸다. 영상( $I$ )의 에지 영상( $E$ )는 아래의 식과 같이 정의된다.

$$E(I) = \|\nabla I\| = \sqrt{\left[\frac{\partial I(x,y)}{\partial x}\right]^2 + \left[\frac{\partial I(x,y)}{\partial y}\right]^2} \quad (10)$$

로고 형판이 반투명 로고이면, 불투명 로고에서의 방법과 같이, 임의의 프레임을 참조 영상으로 삼아 로고의 에지를 추출하는 것이 쉽지 않다. 반투명 로고는 배경에 의존적이어서 변화가 자주 일어난다. 일반적으로 반투명 로고의 내부는 많은 문자 정보를 포함하고 있지 않고, 오히려 단순하기 때문에, LVI로부터 에지 정보를 추출하는 것이 가능하다.

로고 인식을 위하여서는 에지 기반 형판 정합 방법이 사용되었다. 로고 데이터베이스에 있는 로고 에지 형판을 사용하여 입력 에지 영상( $L(i)$ )와의 유사성을 측정하였다. 유사성을 측정하기 위하여서는 아래와 같이 정규 교차 상관 값( $N(m,n)$ )을 사용하였다.

$$N(m,n) = \frac{R(m,n)}{\sqrt{\left[\sum_i \sum_j L(i,j)^2\right] \left[\sum_i \sum_j T(i-m,j-n)^2\right]}} \quad (11)$$

$$R(m,n) = \sum_i \sum_j L(i,j) T_k(i-m,j-n) \quad (12)$$

위의 수식에서  $T_k(i-m,j-n)$ 는 로고 데이터베이스의 로고 에지 영상의  $k$ 번째 형판을 나타내고,  $(i-m,j-n)$ 는 형판 영상에서의 좌표를 나타낸다. 정규 교차-상관 값은 0과 1사이의 범위로 스케일링된다.

## 2.5 로고 전이 검출을 위한 부-비디오 병합

이 절에서는 부-비디오의 병합 방법을 설명한다. 로고 전이를 검출하기 위하여서 이전에 장면 경계 추출 방법을 사용하였다. 그러나 그 분할은 적절한 장면 경계 뿐만 아니라, 불필요한 부분의 장면까지 나누어지게 된

다. 이러한 불필요한 장면은 부-비디오 병합을 통해 제거할 수 있다. 또한, 부-비디오에서의 로고 검출이 실패하였을 경우 전후의 부-비디오를 고려함으로써 보상될 수 있다.

만약에 연속되는 부-비디오들이 같은 로고를 가지고 있다면, 그 둘은 병합된다. 그러나 어떤 비디오는 로고 추출 및 인식을 위한 충분한 프레임을 가지고 있지 않을 수 있다. 만약에 부-비디오의 로고가 추출되고 인식되면, 비디오는 그 로고에 의해서 태그(Tag)된다. 만약에 부-비디오의 로고가 추출이 되지 않거나 인식되지 않으면, 비디오는 전의 태그된 부-비디오나 이후의 부-비디오와 병합되어 로고 인식이 다시 수행된다. 또한, 병합된 비디오에서도 인식이 되지 않으면, 그 부-비디오는 Non-logo로 태그 되고, 일정한 신뢰값(상관값) 이상을 가지는 로고로 인식 되면, 병합된 비디오는 그 로고로 태그 된다. 이러한 과정을 반복하면서 전체 부-비디오의 병합이 이루어진다.

## 3. 실험 및 결과 분석

### 3.1 실험환경

실험을 위해서 300개의 비디오를 수집하였다. 비디오 데이터는 5개의 공중 방송국 채널, 7개의 케이블 채널 그리고 다양한 외국 TV 채널에서 수집되었다. 수집된 방송은 MBC, KBS1, KBS2, SBS, EBS, MTV, Tooniverse, MBN, OCN, Onstyle, Olive, MBC ESPN, Storyon, ABC, CTV, CW, CNN 이다. 비디오 프레임의 크기는 320×320 픽셀로 정규화하였다. 그림 7은 수집된 비디오 데이터의 예를 보여준다. 윗줄의 그림 7(a-c)는 반-투명 로고를 가지고 있는 프레임을 보여주고 있고, 아래 줄의 그림 7(d)는 불투명한 로고를 가지고 있는 프레임들을 보여주고 있다.

### 3.2 실험 결과

로고 인식 실험의 성능을 측정하기 위하여 아래의 식을 이용하였다.

$$\text{Recognition rate}(\%) = \frac{N_r}{N_r + N_{r_f} + N_{r_m}} \times 100 \quad (13)$$

위의 식에서  $N_r$ ,  $N_{r_f}$ ,  $N_{r_m}$ 은 각각 정확하게 인식된 로고의 개수, 잘못 인식된 로고의 개수, 인식하지 못하고 놓친 로고의 개수를 나타낸다. 다양한 임계값( $\tau$ )에 대한 인식 결과를 표 1에서 보여준다.  $\tau$ 는 식 (6)에서 정의된 임계값으로 값이 커질수록 Missing이 줄어드는 반면에, False가 늘어나는 결과를 보여준다. 표 1에서 보여주는 것과 같이 제안된 로고 타입 결정에 의한 인식 방법이 불투명, 반투명 로고들에 대해서 높은 성능을 보여주고 있다. Cozar의 방법이 82.33%의 정확도를 나

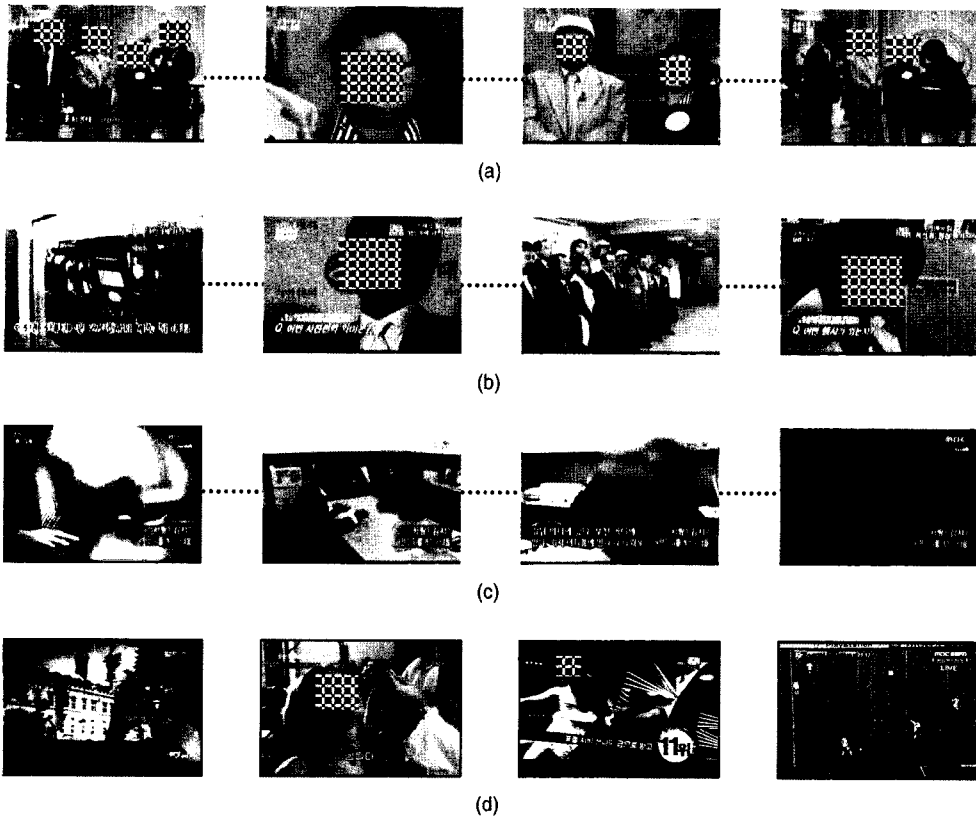


그림 7 실험 비디오 데이터의 예: (a-c) 반투명 로고를 가지는 비디오의 예, (d) 불투명 로고를 가지는 비디오의 예

표 1 로고 인식 결과. ( $\tau$ 는 임계값)

방법		로고타입	전체	Correct	False	Missing	인식률(%)
제안한 방법	$(\tau=0.50)$	불투명	150	142	0	8	82.33
		반투명	150	105	0	45	
	$(\tau=0.75)$	불투명	150	142	0	8	91.67
		반투명	150	133	5	12	
	$(\tau=1.00)$	불투명	150	140	2	8	87.33
		반투명	150	122	17	11	
Cozar 등의 방법[1]		불투명	150	128	8	14	82.33
		반투명	150	119	12	19	

타내는 반면에, 제안된 방법은 최고 91.67%의 정확도를 나타내고 있다. 그림 6에서 보이는 예제 그림과 같이 기존의 방법으로 인식하기 어려웠던 로고들을 로고의 타입에 따른 특징 추출과 인식 방법을 적용함으로써 성공적으로 인식할 수 있었다.

### 3.3 로고 전이 검출

로고 전이 검출 실험을 위해서 각각 19개와 25의 로고를 가지는 2개의 비디오를 데이터로 사용하였다. 로고 전이율을 평가하기 위해 아래와 같이 정의된 Precision

과 Recall을 사용하였다.

$$Precision(\%) = \frac{N_{t_c}}{N_{t_c} + N_{t_f}} \times 100 \quad (14)$$

$$Recall(\%) = \frac{N_{t_c}}{N_t} \times 100 \quad (15)$$

위의 식에서  $N_t$ ,  $N_{t_c}$ ,  $N_{t_f}$ 는 전체 로고의 개수, 로고 전이를 바르게 검출한 개수, 잘못 검출한 개수를 각각 나타낸다.



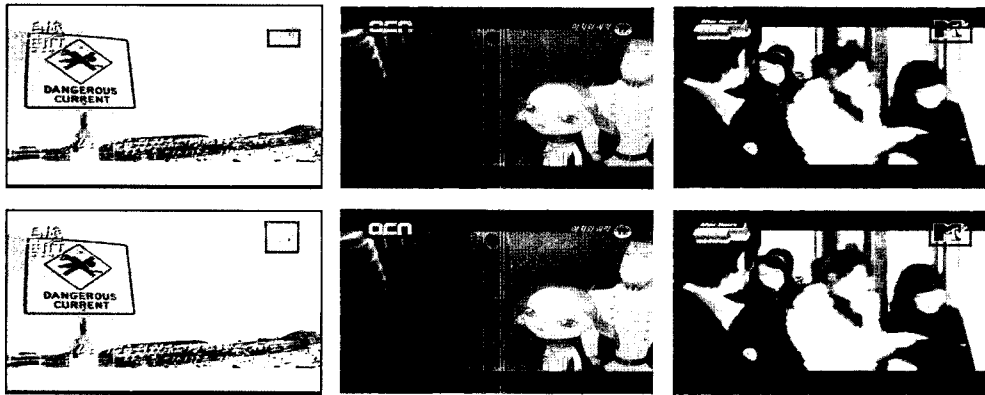
표 2 로고 전이 검출 결과

방법	비디오	전체	Correct	False	Missing	Precision	Recall
제안한 방법	비디오-1	19	19	0	0	100%	100%
	비디오-2	25	24	2	1	92.3%	96%
Cozar의 방법[1]	비디오-1	19	14	0	5	100%	73.7%
	비디오-2	25	18	0	6	100%	72.0%

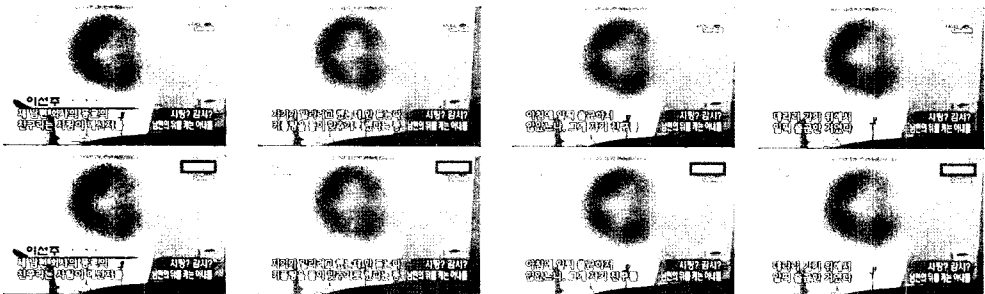
표 2는 로고 전이 검출 결과를 보여준다. 표에서 보는 것과 같이 제안된 방법이 Cozar의 방법보다 좋은 결과를 보여주고 있다. Correct, False, Missing은 각각 로고 전이를 바르게 검출한 경우의 수, 로고의 전이가 없는 곳에서 로고의 전이가 발생하였다고 오검출한 경우의 수(False Alarm), 그리고, 로고 전이가 발생하였는데 검출하지 못한 경우의 수를 나타낸다. 표 2에서, 제안한 방법이 Cozar의 방법보다 훨씬 좋은 Recall을 보여주고 있다. Cozar의 방법에서는 False가 발생하지 않아서 Precision이 높은 결과를 보여주지만, Missing이 많이 발생하여서 Recall이 현저히 떨어짐을 볼 수 있다. 기존의 방법에서는 일정 크기 이상의 동일한 윈도우즈 크기로 비디오를 나누어서 로고를 추출한다. 그렇기 때문에

서로 다른 로고가 하나의 부비디오에 포함되어서 로고를 인식하지 못하는 Missing이 종종 발생하였다. 제안한 방법은 장면 검출을 이용하여 비디오를 나누기 때문에, Missing을 많이 줄일 수 있었다. 반면, 검출된 장면에서의 프레임수가 너무 적게 될 때 정확한 인식을 위한 특징을 충분히 얻지 못해서 False가 발생하였다.

그림 8은 Cozar의 방법(위쪽)과 제안한 방법(아래쪽)에 따른 로고 검출 결과를 보여주고 있다. 그림 8(a)에서 Cozar의 방법은 아래와 위의 검은 배경 때문에 로고의 정확한 위치를 추출하는데 실패 하였다. 제안한 방법에서는 로고 타입에 따른 인식 방법을 사용하여 좋은 결과를 얻을 수 있었다. 그림 8(b)에서는 부-비디오의 배경이 거의 변화가 없어서 인식에 실패하였지만, 제안



(a) 로고 영역의 변화가 거의 없는 비디오의 예



(b) 영상 전체 영역의 변화가 거의 없는 비디오의 예

그림 8 Cozar의 방법(윗줄)과 제안한 방법(아랫줄)에 의한 로고 검출 결과의 비교



그림 9 제안한 방법을 사용한 다양한 비디오에서의 로고 검출 결과

한 방법에서는 부-비디오 병합을 통하여 보상되는 결과를 보여준다. 그림 9에서는 다양한 비디오에서의 실험 결과를 보여준다.

#### 4. 결론 및 향후 연구 방향

사용자에 의해서 편집된 비디오에서는 로고 전이가 여러 번 발생하고, 다른 종류의 로고가 존재하게 된다. 그러므로 로고 타입 결정과 로고 전이 검출의 단계를 통한 효율적인 로고 검출 방법을 제안하였다. 먼저, 장면 검출을 이용하여 비디오는 부-비디오들로 나누어지고, 분할된 각 부-비디오에서 후보 로고 형판이 추출된다. 추출된 각 후보 로고 형판들로부터 로고 타입이 결정되고, 이 로고 타입 결정은 다양한 로고와 다양한 환경에서 강인한 로고 인식을 가능하게 한다. 그 후에, 나누어진 부-비디오들을 병합하기 위하여 병합 방법을 제안하였다. 병합 방법을 통하여 각 부-비디오에서의 로고 검출이 실패하더라도 전후의 부-비디오를 고려해서 보

상할 수 있었다. 제안한 방법은 다양한 TV 비디오에서 실험되어서, 91.67%의 로고 인식 성능을 보여주었고, 다양한 로고를 포함하고 있는 비디오들에서 로고 전이 검출은 평균 95.6%의 Precision과 평균 97.7%의 Recall를 보여주었다. 동일한 실험 데이터에 대하여 Cozar의 방법을 이용하였을 경우, 로고 인식에 있어서는 82.33%의 성능을 보여 주었고, 로고 전이 검출에 있어서는 평균 100%의 Precision과 평균 72.7%의 Recall을 보여줌으로써, 제안한 방법이 더 좋은 결과를 나타냄을 알 수 있었다.

향후에는 로고 전이 검출을 위해 사용된 장면 검출 방법을 개선하고, 보다 범용적인 로고 인식을 위하여 애니메이션 로고 추출 및 인식하는 방법을 연구할 예정이다.

#### 참고 문헌

- [1] J. R. Cozar, N. Guil, J. M. Gonzalez-Linares, E. L. Zapata, and E. Izquierdo, "Logotype Detection to

Support Semantic-Based Video Annotation," *Signal Processing: Image Communication*, Vol. 22, No. 7, pp. 669-679, 2007.

[2] J. R. Cozar, N. Guil, J. M. Gonzalez-Linares, and E. L. Zapata, "Video Cataloging Based on Robust Logotype Detection," *Proc. 2006 IEEE International Conference on Image Processing*, Atlanta, USA, pp. 3217-3220, October 2006.

[3] W. Q. Yan, J. Wang, and M. Kankanalli, "Automatic Video Logo Detection and Removal," *Multimedia Systems*, Vol. 10, No. 5, pp. 379-391, 2005.

[4] H. Fang, J. Jiang, and Y. Feng, "A Fuzzy Logic Approach for Detection of Video Shot Boundaries," *Pattern Recognition*, Vol. 39, No. 11, pp. 2092-2100, 2006.

[5] J. Wang, Q. Liu, L. Duan, H. Lu, and C. Xu, "Automatic TV Logo Detection, Tracking and Removal in Broadcast," *Lecture Notes in Computer Science: Advances in Multimedia Modeling*, Vol. 4352, pp. 63-72, 2007.

[6] K. Meisinger, T. Troeger, M. Zeller, and A. Kaup, "Automatic TV Logo Removal using Statistical Logo Detection and Frequency Selective Inpainting," *Proc. 13th European Signal Processing Conference*, Antalya, Turkey, September 2005.

[7] A. Albiol, M. J. Fulla, A. Albiol, and L. Torres, "Commercials Detection Using HMMs," *Proc. IEEE International Workshop on Image Analysis for Multimedia Interactive Services*, Lisboa, Portugal, 2004.

[8] J. H. Yeh, J. C. Chen, J. H. Kuo, and J. L. Wu, "TV Commercial Detection in News Program Videos," *Proc. IEEE International Symposium on Circuits and Systems*, Kobe, Japan, Vol. 5, pp. 4594-4597, May 2005.

[9] A. Albiol, M. J. Fulla, A. Albiol, and L. Torres, "Detection of TV Commercials," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Montreal, Canada, Vol. 3, pp. 541-544, May 2004.

[10] X. Tong, H. Lu, Q. Liu, and H. Jin, "Replay Detection in Broadcasting Sports Video," *Proc. 3rd International Conference on Image and Graphics*, Hong Kong, China, pp. 337-340, December 2004.

[11] B. Seeber, N. Yager, and A. Amin, "Real-time Detection of Semi-transparent Watermarks in Decompressed Video," *Proc. IEEE Workshop on Applications of Computer Vision*, Texas, USA, pp. 49-54, February 2007.

[12] S. Duffner and C. Garcia, "A Neural Scheme for Robust Detection of Transparent Logos in TV Programs," *Lecture Notes in Computer Science: Artificial Neural Networks*, Vol. 4132, pp. 14-23, 2006.

[13] H. Bai, W. Hu, T. Wang, X. Tong, C. Liu, and Y.

Zhang, "A Novel Sports Video Logo Detector Based on Motion Analysis," *Lecture Notes in Computer Science: Neural Information Processing*, Vol. 4233, pp. 448-457, 2006.

[14] L. Boussaid, A. Mtibaa, M. Abid, and M. Painsdavoine, "A Real-Time Shot Cut Detector: Hardware Implementation," *Computer Standards and Interfaces*, Vol. 29, No. 3, pp. 335-342, 2007.



노 명 철

2001년 강원대학교 컴퓨터공학과(학사)  
2003년 고려대학교 컴퓨터학과(석사). 2008  
년 고려대학교 컴퓨터학과(박사). 2004  
년~2005년 University of Surrey 방문  
연구원. 2008년~현재 Carnegie Mellon  
University 관심분야는 컴퓨터 시각, 패

틴인식, 뇌공학 등



강 승 연

2006년 한신대학교 정보통신학과(학사)  
2008년 고려대학교 컴퓨터학과(석사). 관  
심분야는 컴퓨터 시각, 패턴인식, 영상처  
리 등



이 성 환

1984년 서울대학교 계산통계학과(학사)  
1986년 KAIST 전산학과(석사). 1989년  
KAIST 전산학과(박사). 1989년~1995년  
충북대학교 컴퓨터학과 전임강사, 조교  
수. 2001년~2002년 MIT Brain and  
Cognitive Sciences 학과 방문교수. 1995

년~현재 고려대학교 정보통신대학 정교수. 1997년~현재  
고려대학교 인공시각연구센터 소장. 관심분야는 패턴인식,  
컴퓨터시각, 뇌공학 등