

論文2003-40SP-1-7

새로운 결합척도를 이용한 동영상 분할

(Video Segmentation Using New Combined Measure)

崔在覺*, 李時雄**, 南在悅***

(Jae Gark Choi, Si-Woong Lee, and Jae-Yeal Nam)

요 약

본 논문에서는 분할기반 영상 부호화를 위한 새로운 영상 분할 알고리즘을 제안한다. 제안된 방법은 움직임과 밝기 정보에 기반한 새로운 유사성 척도를 사용한다. 그리고 하나의 분할 단계 내에 밝기와 움직임 정보가 함께 결합된다. 영상 분할은 분수령 알고리즘에 기반한 영역 확장법을 통해 이루어지며, 연속된 프레임에 대한 분할은 분할결과가 시간축으로 일관성을 유지하도록 추적방법을 통해 이루어진다. 모의실험결과, 제안된 방법이 통계적 척도만을 사용한 방법과는 달리, 물체의 경계를 결정하는데 효과적임을 보였다.

Abstract

A new video segmentation algorithm for segmentation-based video coding is proposed. The method uses a new criterion based on similarities in both motion and brightness. Brightness and motion information are incorporated in a single segmentation procedure. The actual segmentation is accomplished using a region-growing technique based on the watershed algorithm. In addition, a tracking technique is used in subsequent frames to achieve a coherent segmentation through time. Simulation results show that the proposed method is effective in determining object boundaries not easily found using the statistic criterion alone.

Keyword : morphological segmentation, motion estimation, object-based video coding, object tracking

I. 서 론

기존의 영상 분할 방법은 분할시 사용하는 척도(criterion)에 따라서 크게 두개의 부류로 나누어진다. 그중 하나는 움직임의 균일성을 분할의 척도로 삼는

움직임기반 물체 분할법들로서 [1][2][3]등이 이에 해당한다. 이 기법들은 가정된 움직임 모형에 근거하여, 동일한 움직임으로 표현되는 영역으로 영상을 분할하는 것으로, 움직임 추정의 부정확도에 의해서 근본적인 한계를 지닌다. 즉 움직임 영역의 분할이 정확하기 위해서는 우선 움직임이 정확히 추정이 되어야 하는데, 기존의 움직임 추정기법들은 저주파 영역이나 움직임 경계에서 오류가 발생한다. 따라서 잘못 추정된 움직임 정보에 근거하여 분할을 행하면 이러한 부분에서 부정확한 분할 결과를 초래하며, 특히 추출된 움직임 경계가 매우 거칠어진다.

다른 부류는 영상 밝기값의 연속성을 분할 척도로 하는 영역 분할법이다. 영역 분할법으로 CLRG(centroid linkage region growing), MST(minimum spanning tree), RAG(region adjacency graph)등의 알고리즘이

* 正會員, 東義大學校 컴퓨터工學科
(Dongyeui University, Division of Computer Application Eng.)

** 正會員, 한밭大學校 情報通信·컴퓨터工學部
(Taejon National University of Technology, School of Information Communication and Computer Eng.)

*** 正會員, 啓明大學校 컴퓨터工學科
(Keimyung University, Dept. of Computer Eng.)

接受日字:2002年6月4日, 수정완료일:2002年11月15日

있다. 특히 요즘에는 수리형태학(mathematical morphology)에 근거한 영역 분할법이 활발히 연구되고 있다. 이는 수리형태학이 영역의 크기, 모양 및 대조 등의 영역 분할에 매우 적합한 성질들에 바탕을 두기 때문이다. 이 기법들은 대부분 계층구조로 구성되며, 각 계층에서는 영상 단순화(image simplification), 마커 추출(marker extraction), 경계선 결정(decision) 및 화질 검사(quality test)의 네 과정으로 이루어져 있다. 각 계층에서 분할된 영역은 영역 경계 정보와 내용(texture) 정보로 표현되거나^[4], 움직임 보상을 행하는 경우에는 영역 경계 정보, 움직임 정보 및 오차 신호 정보로서 표현된다^[5]. 참고문헌 [4][6][7]의 방식은 부호화시에 시간 방향의 중복성을 고려하지 않는다는 점에서 단점을 가진다. 이를 극복하기 위하여 참고문헌 [5]에서는 움직임 추정 및 보상 방식을 이용하는 구조가 제시되었다. 하지만, 영역 분할의 척도로 밝기값만을 사용하기 때문에 분할된 영역의 경계가 움직임의 경계와는 일치하지 않으며, 따라서 움직임 추정 및 보상이 부정확하게 된다. 이것은, 인접 영역의 밝기가 서로 비슷하다면 비록 다른 움직임을 갖더라도 분할시에 이를 구분할 수 없기 때문이다. 한편, 참고논문 [8]에서는 정확한 비디오 객체 분할을 위하여 사용자의 조력(user's assistance)을 이용하는 반자동 영상분할 방법이 제안되었다.

본 논문에서 제안되는 영상분할방법은 다음의 두 조건을 기준으로 고안되었다.

- 움직임을 이용하여 영상을 분할함으로써 영상 부호화 적용시 움직임 보상 효율의 증대
- 연속된 영상간의 분할 상관성을 최대화시킴으로써 윤곽선 부호화 효율의 증대

조건 1은 기존의 영역 분할법들이 영상 분할시에 영상의 밝기값만을 분할 척도로 사용함으로써, 분할된 영역이 움직임의 관점에서는 균일성이 보장되지 못한다는 단점을 극복하기 위한 것이다. 그리고, 조건 2는 분할 기반 부호화의 관점이 윤곽선 정보의 감축에 있음을 고려하여 설정한 것이다. 즉, 윤곽선 정보가 전체 정보량 중에서 매우 큰 비중을 차지하며 이를 효과적으로 압축하기 위해 윤곽선 신호의 프레임간 부호화가 요구되기 때문에 프레임간에 높은 상관성을 갖는 영상 분할이 필수적이다. 따라서 본 논문에서는 이러한 기준을 만족하기 위해 영역 분할의 척도로 움직임의 연속

성과 밝기값의 연속성을 모두 고려하는 분할법을 제시하고자한다. 그리고, 연속된 영상사이에 영상 분할의 상관성이 최대가 되도록 이전 영상의 분할 결과를 움직임 정보를 이용하여 현재 영상에 투영하여 현재 영상을 분할하는 방법을 사용한다.

본 논문의 구성은 다음과 같다. 우선 2절에서 수리형태학을 이용하는 영역 분할 방식들과 그들의 문제점을 살펴본 후, 3절에서 제안된 분할 기법을 설명하기로 한다. 4절에서 모의 실험을 통하여 제안 방식의 성능을 검토하고, 5절에서 결론을 맺는다.

II. 수리형태학에 근거한 제 분할법 및 문제점

1. 수리형태학적 영상분할의 도구

본 절에서는 우선 수리형태학적 영상분할의 도구들(morphological tool)의 특성에 대하여 살펴보기로 한다. 기본적으로 수리형태학을 이용한 영상 분할은 모포로지컬 필터를 이용한 영상 단순화와 분수령 알고리즘(watershed algorithm)을 이용한 경계선 결정의 과정으로 구성된다.

먼저 모포로지컬 필터에 대해 살펴본다. 다양한 모포로지컬 필터들은 아래에서 정의되는 erosion과 dilation의 두 기본적인 변환에 근거하여 구성된다.

$$\text{Erosion} : \varepsilon_n(f)(x) = \text{Min}\{f(x+y), y \in M_n\} \quad (1)$$

$$\text{Dilation} : \delta_n(f)(x) = \text{Max}\{f(x-y), y \in M_n\} \quad (2)$$

여기서 $f(x)$ 는 입력신호이며 M_n 은 크기가 n 인 평면 structuring element이다. 그리고 크기가 1인 geodesic dilation 및 erosion은 다음처럼 정의된다.

$$\text{Geodesic dilation of size 1} : \delta^{(1)}(f, r) = \text{Min}\{\delta_1(f), r\} \quad (3)$$

$$\text{Geodesic erosion of size 1} : \varepsilon^{(1)}(f, r) = \text{Max}\{\varepsilon_1(f), r\} \quad (4)$$

크기가 1인 geodesic dilation의 출력값은 원 신호인 f 의 dilation된 값과 기준신호 r 중에서 최소치로 정의되며, geodesic erosion은 geodesic dilation과 duality 관계에 있기 때문에 다음처럼 표시될 수 있다.

$$\varepsilon^{(1)}(f, r) = -\delta^{(1)}(-f, -r) \quad (5)$$

크기가 n 인 geodesic dilation이나 erosion은 크기가 1인 geodesic 변환의 n 차 반복적 구성으로 구현된다. 특히, 무한대 크기의 geodesic dilation(erosion)을 dilation(erosion)에 의한 재생이라 부르며, 다음 식으로 정의된다.

Reconstruction by dialation :

$$\gamma^{(rec)}(f, r) = \delta^{(\infty)}(f, r) = \dots \delta^{(1)}(\dots \delta^{(1)}(f, r) \dots, r) \quad (6)$$

Reconstruction by erosion :

$$\phi^{(rec)}(f, r) = \varepsilon^{(\infty)}(f, r) = \dots \varepsilon^{(1)}(\dots \varepsilon^{(1)}(f, r) \dots, r) \quad (7)$$

이상에서 설명된 dilation과 erosion의 두 기본적인 연산자를 이용하여 각 모포로지컬 필터를 구성한다. 우선 모포로지컬 opening과 closing은 아래와 같이 정의된다.

$$\text{Morphological opening : } \gamma_n(f) = \delta_n(\varepsilon_n(f)) = \delta_n \varepsilon_n \quad (8)$$

$$\text{Morphological closing : } \phi_n(f) = \varepsilon_n(\delta_n(f)) = \varepsilon_n \delta_n \quad (9)$$

Morphological opening 필터는 원 신호에서 structuring element보다 크기가 작은 밝은 부분들을 제거하고, 반대로 morphological closing 필터는 어두운 부분들을 제거함으로써 영상을 단순화한다. 여기서 밝고 어둡다는 것은 모두 주위의 신호값에 대한 상대적인 표현이다. 즉 식 (1)과 식 (2)에서 보듯이 morphological opening의 경우에 erosion 과정에서 structuring element보다 작은 밝은 부분들은 모두 주위의 보다 어두운 신호값으로 대체되기 때문에, 그 부분의 밝기 정보가 제거된다. 그런데, morphological opening이나 dilation은 모두 비대칭적(non-symmetrical) 필터이다. 즉, 밝고 어두운 것중 어느 한쪽만을 제거할 뿐이다. 따라서, 대칭적인 특성을 갖는 필터를 구성하려면 open_close ($\gamma_n(\phi_n(f))$) 필터나 close_open ($\phi_n(\gamma_n(f))$) 필터를 사용한다.

이 필터들은 모두 영상분할 전단계에서 영상 단순화를 위한 목적으로 사용될 수 있는 비선형 필터이다. 그러나 morphological opening이나 closing 필터를 사용하여 영상을 단순화하면 윤곽선 정보가 제대로 보존되지 않고 손상된다는 것이 알려져 있다^[4]. 윤곽선 정보가 보존되지 않으면 정확한 경계선 검출이 불가능하기 때문에, 윤곽선을 유지시켜주는 필터의 사용이 필수적이다. 이러한 필터로서 재생에 의한 필터가 있다. 재생에 의한 opening 및 closing 필터는 다음처럼 정의된다.

$$\text{Opening(closing) by reconstruction of erosion (dilation) : } \gamma^{(rec)}(\varepsilon_n(f), f), \phi^{(rec)}(\delta_n(f), f) \quad (10)$$

재생에 의한 opening인 경우에 단순화는 structuring element보다 작은 밝은 부분을 제거하는 erosion 과정에서 수행된다. 그리고 제거되지 않은 큰 영역의 경계선은 재생과정에서 완전하게 복구된다.

2. 수리형태학적 도구를 이용한 동영상 분할 방법

동영상에서 연속된 영상들의 각 프레임을 독립적으로 분할하고 이렇게 분할된 결과를 이용하여 분할기반 영상 부호화를 행할 경우, 다음과 같은 문제가 있다. 이웃한 영상들은 서로간에 많은 상관성을 포함함에도 불구하고, 각 영상이 독립적으로 분할되었기 때문에 이웃한 영상의 분할 결과들은 상관성이 없이 분할되어 나타난다. 따라서 분할된 영역의 모양정보 즉, 윤곽선 정보를 부호화 하는데 많은 정보량이 필요함으로 윤곽선 정보에 많은 비트가 할당되어 부호화 효율이 감소된다.

이러한 문제를 해결하기 위해 참고문헌 [4]에서는 동영상을 3차원 공간상의 함수로 생각하여 동영상을 3차원 영역으로 분할한다. 3차원 영역으로 영상을 분할하면 시간축으로 안정된 분할 결과를 얻을 수 있을 뿐 아니라 영역 정합문제를 쉽게 해결할 수 있다. 참고문헌 [4]에서는 동영상을 시간축이 포함된 3차원 함수로 생각하여 모포로지컬 필터를 3차원 동영상에 적용하여 영상을 단순화하고, 단순화된 동영상에서 밝기가 일정한 부분을 마커로 검출한 뒤, 분수령 알고리즘을 3차원적으로 적용하여 동영상을 분할하게 된다.

동영상 분할 과정은 top-down 방식의 계층적 구조이며 각 계층은 단순화, 마커 추출, 경계선 결정, 화질 추정,의 네단계로 구성된다. 자세한 내용은 다음과 같다.

- 단순화 단계: 단순화는 입력 신호에서 현 계층에서의 분할에 필요한 정보외의 세부적인 신호를 제거하는 과정이다. Coding residue가 입력 신호이며, 단순화를 위한 필터로는 재생에 의한 open_close 나 close_open 필터가 주로 사용된다. Coding residue란 이전 계층에서 복호된 영상과 원 영상과의 차신호를 뜻한다.
- 마커 추출 단계: 마커 추출은 단순화된 신호에서 국부적으로 균일한 밝기값을 갖는 3차원 블록을 찾는 과정이다. 이 마커들은 단지 영역의 내부를 표시할 뿐이

며, 정확한 영역간의 경계는 결정하지 않는다.

- 경계선 결정 단계: 결정단계에서는 단순화된 신호와 마커정보를 입력으로 하여 영역간의 경계를 watershed algorithm으로 결정한다. 이 단계에서 현 계층의 영역 분할 결과가 나오며 각 분할된 영역은 그 영역의 값으로 labelling된다.

- 화질 추정 단계: 3차원으로 분할된 각 영역은 부호화 과정을 통하여 영역의 정보를 전송하게 된다. 이 정보로부터 수신될 영상을 재구성하고 이 영상과 원 영상과의 차이(coding residue)로부터 하부계층의 분할 여부를 판단한다. 오차신호가 크면 잘 표현되지 못한 영역이 존재하는 것이므로 다음 계층으로 내려간다.

이 기법에서는 분할된 영역의 정보를 표현하기 위해 영역의 경계선 정보와 영역의 컬러 정보를 전송한다. 영역의 컬러 신호는 기본적으로 저차원의 다항식으로 표현되며, 근사화 오차를 BTC 기법으로 전송한다. 이 분할 방법의 문제점을 분석하면 다음과 같다.

- 이 방법은 기본적으로 3차원 분할기법이므로 3차원 필터링을 위하여 많은 메모리와 계산 시간이 요구되고, 프레임의 지연이 발생한다. 특히 재생에 의한 open_close 나 close_open 필터가 사용되기 때문에 많은 계산량이 필요하다.

- 3차원 분할시 움직임 정보는 전혀 이용하지 않고 밝기 정보 또는 컬러 정보만을 이용하여 분할하기 때문에 영상분할의 정확도가 줄어든다.

- 움직임 정보를 사용하지 않고 영상 분할을 하기 때문에, 부호화시 부호화할 정보는 컬러정보와 윤곽선 정보로 표현되고 따라서 움직임을 이용한 부호화에 비해

부호화 효율이 저하된다.

III. 제안된 영상 분할 기법

제안된 동영상 분할 방법은 두 개의 분할 모드를 갖는다. 프레임내(intra-frame) 분할 모드는 동영상의 첫 번째 영상에 대해 적용되는 알고리즘이며, 한 장의 영상만을 사용하여 영상을 움직임 물체별로 분할하게 된다. 동영상의 첫 번째 영상이 프레임내 분할 모드를 사용하여 분할되었다면, 동영상의 다음 프레임들은 프레임간 분할 모드를 이용하여 영상을 분할하게 된다. 프레임간(inter-frame) 분할 모드는 이전 프레임의 분할 결과를 움직임 정보를 이용하여 현재 프레임에 투영함으로써 현재 프레임을 분할하는 기법이다. 이 두 모드의 분할과정을 다음 소절에서 자세히 기술한다.

1. 결합척도에 의한 프레임내 분할

제안된 프레임내 영상분할방법은 <그림 1>에 나타난 과정을 거쳐서 분할된다. 블록도에서 알 수 있듯이 마커 추출과 경계선 결정 과정에서 제안 기법은 영상의 밝기값과 움직임 정보를 모두 고려한다. 우선 단순화된 현재 영상으로부터 밝기 마커(intensity marker)를 추출하고, 두 연속된 영상간의 움직임 벡터장에서 움직임 마커(motion marker)를 추출한다. 밝기 마커는 균일한 밝기값을 갖는 부분이며, 분할될 영역의 내부에 위치한다. Motion 마커는 균일한 움직임을 갖는 부분으로서, 이동체의 내부를 표현하게 된다. 다음에는 이 두 종류의 마커 정보로부터 최종적인 결합 마커(combined marker) 들을 추출한다. 결합 마커는 움직임과 밝기값이 모두 균일한 영역으로, 경계선 결정 과정에서 seed

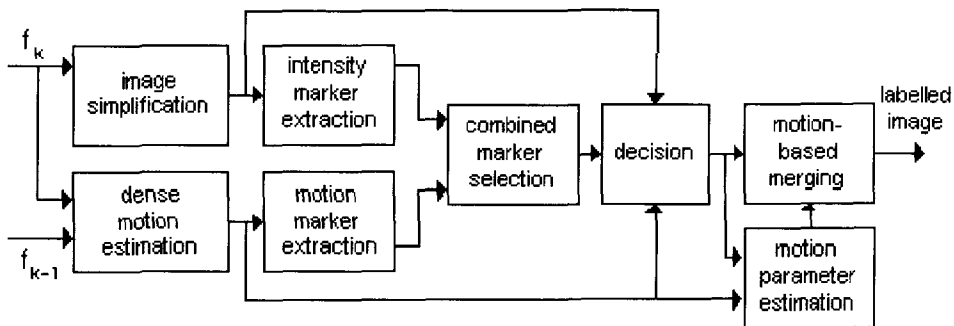


그림 1. 프레임내 분할방법의 블록도

Fig. 1. Block diagram of intra-frame segmentation method.

영역으로 사용된다. 경계선 결정을 통해 영역의 경계선이 결정되면 움직임을 근거로 하여 각 영역을 병합함으로써 최종적인 분할 결과가 얻어진다.

(1) 영상 단순화

영상의 단순화를 위한 필터로는 앞 절에서 설명된 재생에 의한 open_close 필터를 사용한다. 주파수 대역 제한을 목적으로 일반적으로 사용되는 선형 필터들은 영역 내부에 있는 고주파를 제거함과 동시에 영역 경계를 뭉개기 때문에 영역 분할을 위한 전처리 필터로는 부적당하다. 반면, 재생에 의한 open_close 필터는 structuring element의 내부에 들어오는 작은 영역들은 모두 제거하면서, 큰 영역들의 경계를 그대로 유지시켜 주므로 영역 분할용 단순화 필터에 적합한 우수한 특성을 지닌다.

(2) 밝기 마커 추출

밝기 마커란 균일한 밝기값으로 연결된 영역으로서, 균일한 영역을 검출하여 그 부분을 label로 표시함으로써 추출된다. 균일하지 않거나 두 마커의 중간 부분은 불확실 영역이므로 0으로 표시한다. 단순화된 영상에서 밝기 마커를 추출하는 방법은 영상의 단순화 특징에 따라서 두 가지로 나눌 수 있다.

- 평탄한 영역에서의 마커추출 : 이 방식은 단순히 동일한 밝기값으로 연결되어 있는 영역을 찾아서 labeling 하는 방법이다. 연결된 영역을 찾는 방법은 queue구조를 이용하는 방법과 blob coloring등의 여러가지가 있다. 추출된 평탄 영역 중에서 크기가 주어진 임계값 이상인 영역만을 마커로서 labeling된다.

- 대조 영역에서의 마커추출 : 밝기값의 대조에 근거한 마커추출은 원 영상(f)과 $f, \gamma_n \phi_n, \phi_n \gamma_n$ 의 중앙값과의 차이에서 얻어진다. $Open_close(\phi_n \gamma_n)$ 와 $close_open(\gamma_n \phi_n)$ 필터들은 각각 상위 및 하위 포락선을 검출하기 때문에 식 (11)의 과정에서 정확히 마커영역이 검출된다.

Contrast extraction : $|f - \text{Min}\{\gamma_n \phi_n, \text{Max}\{\phi_n \gamma_n, f\}\}|$ (11)

(3) 움직임 마커 추출

움직임 마커는 가정된 움직임 모형하에서 균일한 움직임을 갖는 영역이다. 그러므로 움직임 마커는 움직임 벡터장을 균일한 움직임 영역으로 분할함으로써 추출되며, 이를 위한 움직임 벡터장 분할 기법으로 다단계

순환적 영역 분할법을 제안한다.

본 논문에서 사용될 움직임 모형은 평면 강체의 3차원 움직임이 평행투영되어 영상평면상에 나타날 때 유도되는 6계수 모형을 사용한다. 즉,

$$\begin{aligned} V_x(x, y) &= a_1 + a_2x + a_3y \\ V_y(x, y) &= a_4 + a_5x + a_6y \end{aligned} \quad (12)$$

여기서 $V_x(x, y)$ 및 $V_y(x, y)$ 는 움직임 벡터의 수평, 수직 방향의 성분이다.

<그림 2>는 벡터장 분할기의 순서도를 나타낸다. 다단계 순환적 영역수정법을 이용한 벡터장 분할 기법은 매 단계마다 한개씩의 이동체 영역을 분할하는 순차적 과정으로 구성되며, 각 단계는 두개의 세부과정으로 나누어진다. 우선 첫단계에서는 이동체의 움직임 파라미터에 대한 초기값을 추정하게 될 seed block 위치를 검출한다. 두번째 단계에서는 Seed block에서 추정된 움직임 파라미터의 초기값을 순환적 영역 수정법을 통하여 실제의 움직임 파라미터로 수렴시키고 이를 이용하여 이동체 영역(움직임 마커)을 분할한다.

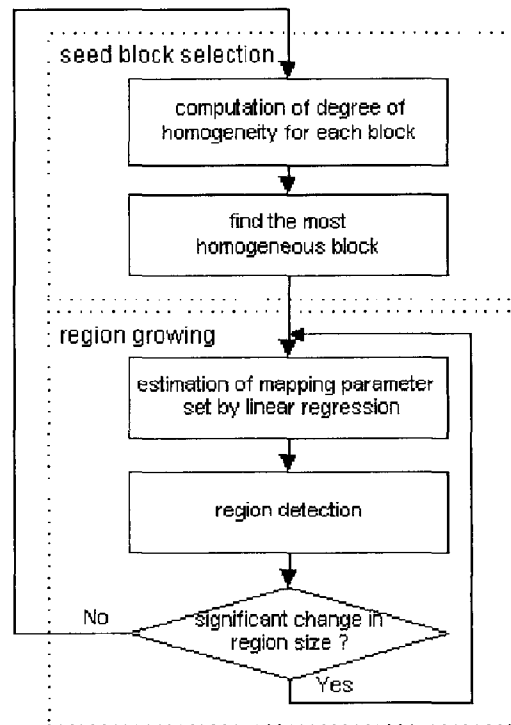


그림 2. 벡터장 분할기의 순서도
Fig. 2. Flow chart of motion vector field segmentation scheme.

Seed block은 이동체 영역의 내부에 위치해야한다. 그러므로 영역의 경계를 모르고 있는 경우에 가장 균일한(homogeneous) 벡터들을 포함하는 블록을 seed block으로 판정하는 것이 가장 적합하다. Seed block은 사각형의 미끄럼 창(sliding window)을 벡터장 내부로 미끄러뜨리며 각 위치에서 창내부의 벡터들의 균일성 정도(degree of homogeneity)를 계산한 후, 이를 비교하여 최대값을 가지는 위치를 선택함으로써 검파한다. 균일성 정도는 주어진 벡터들과 그것들로부터 추정된 움직임 파라미터에 의한 예측치간의 오차를 측정하여 계산한다. 즉,

$$D(m, n) = \frac{1}{E_{mn}(a_1^*, \dots, a_6^*)} \quad (13)$$

여기서,

$$E_{mn}(a_1, \dots, a_6) = \sqrt{\sum_{x=m}^{m+W-1} \sum_{y=n}^{n+W-1} \epsilon(x, y; a_1, \dots, a_6)} \quad (14)$$

$$\epsilon(x, y; a_1, \dots, a_6) = (V_x(x, y) - \widehat{V}_x(x, y))^2 + (V_y(x, y) - \widehat{V}_y(x, y))^2 \quad (15)$$

$$\widehat{V}_x(x, y) = a_1 + a_2x + a_3y \quad (16)$$

$$\widehat{V}_y(x, y) = a_4 + a_5x + a_6y \quad (17)$$

이며, $\{a_1^*, a_2^*, \dots, a_6^*\}$ 는 창 내부의 벡터들에 최적인 움직임 파라미터로서 최소평균오차 척도에 의해 계산되어진다.

Seed block이 검파되면 이 블록으로부터 움직임 파라미터의 초기값 및 초기 영역이 얻어진다. 이 초기값을 실제의 파라미터값으로 수정하기 위해서 영역수정법을 통한 수렴과정을 거친다. 즉, 초기 영역에서 새로 파라미터를 추정하고 그에 따른 영역을 추출한다. 이때 새로 추출된 영역은 새로운 움직임 파라미터에 대해 이전 영역내에서 큰 오차를 갖는 벡터들은 제외시키는 반면, 이전 영역에 포함되지 못한 벡터들 중에서 임계값보다 작은 오차를 갖는 벡터들을 병합함으로써 구성된다. 만약 추출된 영역이 초기 영역에 비해 많이 변했다면 이 순환과정을 수렴할 때까지 반복한다. 반복여부에 대한 판정식은 아래와 같다.

$$\frac{|A(S_i^k) - A(S_i^{k-1})|}{A(S_i^{k-1})} \leq T_s \quad (18)$$

여기서 $A(S_i^k)$ 는 k 번째 단계(k 번째의 이동체)에서 i 번째 순환과정에서 추출된 영역의 크기를 나타낸다. 반

복과정이 끝나면 최종적으로 얻어진 영역 및 움직임 파라미터가 이 단계에서 검출된 움직임 마커 및 그 영역의 움직임 파라미터로 등록되고, 새로운 마커의 검출을 위해 다음 단계로 넘어간다. 만약, 임계값보다 큰 균일한 영역이 발견되지 않으면 모든 움직임 마커는 이미 추출되고 noisy한 영역만이 남아있는 것이므로 이 부분들은 불확정 영역으로 표시하고 과정을 끝낸다.

(4) 결합 마커 추출

결합 마커 영역은 움직임 및 밝기값 모두가 균일한 연결 영역이다. 밝기 마커 및 움직임 마커 정보로부터 이를 추출하는 과정은 다음과 같다. 우선, 각각의 밝기 마커 영역에 대해서 움직임 마커 영역의 분포를 살펴본다. 만약, 밝기 마커 내부에 움직임 마커가 존재하지 않으면, 이것은 움직임이 균일하지 않은 영역임을 나타내므로 마커에서 제거한다. 그리고, 하나의 밝기 마커 영역내에 다수개의 움직임 마커가 걸쳐있다면 그 마커는 내부 화소를 공유하는 움직임 마커의 개수에 의해 나누어진다. 즉, 같은 밝기 마커 내부라 하더라도 움직임 마커의 label이 다르면 다른 이동체 영역이기 때문에, 그 각각을 다른 결합 마커로 추출하는 것이다. 이때, 추출된 결합 마커의 영역 크기가 임계값 이하이면 이를 제외한다.

(5) 경계선 결정

추출된 마커를 중심으로 영역간의 경계를 구분하는 경계선 결정 과정은 참고문헌 [4]에서 설명된 변형된 분수령 알고리즘(modified watershed)을 사용하였다. 그리고 본 논문에서는 분수령 알고리즘에서 사용되는 유사성 척도를 위해 식 (19)에 정의되는 결합척도를 제안한다. 식 (19)의 유사성 척도는 밝기 정보와 움직임 정보를 결합함으로써 밝기가 유사하지만 움직임이 서로 다른 영역을 쉽게 구별하여 정확한 영역 경계 결정을 가능하게 한다.

$$P(x, y) = \frac{\alpha}{P_{i(x, y)}} + \frac{1-\alpha}{P_{m(x, y)}} \quad (19)$$

여기서, $P(x, y)$ 는 (x, y) 위치의 화소가 인접한 영역과의 유사성을 나타내며 α 는 0과 1 사이의 가중치이며, $P_{i(x, y)}$ 및 $P_{m(x, y)}$ 는 각각 밝기값과 움직임의 연속성에 의한 유사성으로서 다음처럼 계산된다.

$$P_{i(x, y)} = \left| \frac{\sum_{(i, j) \in R} I(i, j)}{\sum_{(i, j) \in R} 1} - I(x, y) \right| \quad (20)$$

$$P_{m(x,y)} = \left[\frac{k \epsilon(x,y, \{a_R\})}{step} \right]_{step} \quad (21)$$

여기서, $I(x,y)$ 는 (x,y) 화소의 밝기값을, R 은 (x,y) 에 이웃하는 영역을 나타낸다. 그리고, $[x]$ 는 x 를 넘지않는 최대 정수이다.

식 (19)의 α 를 조정함으로써 두 종류의 연속성에 대한 중요도를 조절할 수 있다. 즉 α 의 값이 크면 밝기 값의 연속성이, 작으면 움직임의 연속성이 분할 척도의 주도적인 역할을 하게된다. 식 (21)의 k 는 움직임 유사성의 값이 밝기 유사성의 값과 비슷한 범위를 갖게 하기 위해 도입되었다. 그리고, step은 밝기 유사성의 값에 대한 양자화 크기인데, 이는 움직임 오차의 작은 변화에 의해 윤곽선이 흐트러지는 것을 방지하는 역할을 한다.

(6) 움직임 기준 영역 병합

앞 소절에서의 영상 분할은 결합 유사성 척도를 사용하여 밝기와 움직임이 동시에 균일한 영역으로 영상을 분할한다. 따라서 밝기는 다르지만 움직임 파라미터는 같은 영역이 존재할 수 있다. 그러므로 부호화 관점에서 이러한 중복된 영역의 제거가 필요하다. 이것이 움직임 기준 영역 병합에 해당하며 영상 분할을 단순화하여 분할된 영역의 윤곽선 및 움직임 파라미터 정보를 크게 줄인다. 따라서 분할된 영역들이 움직임이 유사하고 서로 이웃해 있는 경우 이 영역들에 대해 병합을 행하게 되는데 이때 움직임의 유사성은 식 (12)에 설명된 움직임 모형을 사용한다. 그리고 움직임 기준 영역 병합에 사용된 영역 병합 방법은 참고문헌 [6]에 기술된 방법을 사용하였다.

2. 분할의 시간축 연결을 위한 프레임간 분할

영상 시퀀스의 두번째 프레임부터는, 이미 이전 영상의 분할 결과에 현재 영상의 영상 분할에 대한 많은 정보가 담겨 있기 때문에 시간축 연결 기법을 통하여 영상을 분할한다. 시간축 연결 기법에 의한 영상 분할의 블록도가 <그림 3>에 나와 있다. 이전 영상의 분할 결과를 이용하여 현재 영상을 분할하는 분할결과와 시간축 연결기법은 기본적으로 이전 영상에서 분할된 영역 또는 물체들을 움직임 정보를 이용하여 현재 영상으로 추적(tracking)함으로써 이루어진다. 이전 프레임의 역방향 움직임으로부터 순방향 움직임을 추정한다. 영상의 움직임이 불규칙적으로 변하지 않으면 이 움직임이 현재 프레임의 순방향 움직임에 매우 근접한 값

이라고 생각할 수 있으므로 이 값을 이용하여 이전의 분할 결과에 순방향 움직임 보상을 행한다. 그 다음, 예측된 각 분할 영역의 경계에서 불확실 영역을 추출하고 이 불확실 영역에 대하여 분수령 알고리즘을 적용함으로써 현재 프레임의 정확한 경계선을 결정한다.

(1) 순방향 움직임 추정 및 이전 분할의 순방향 보상
현재 영상의 (x,y) 에 대응하는 이전 프레임의 좌표를 (x',y') 라고 하면 식 (12)로부터 다음의 식이 나온다.

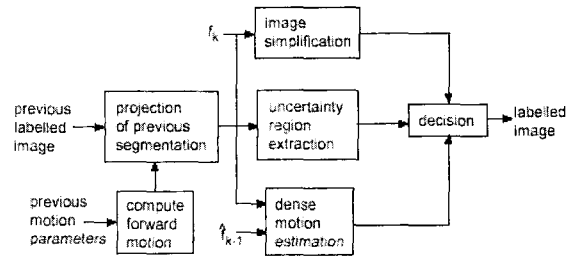


그림 3. 시간축 연결 기법에 의한 프레임간 영상 분할
Fig. 3. Inter-frame segmentation by temporal connection.

$$\begin{aligned} x &= -a_1 + (1-a_2)x - a_3y \equiv a_1 + a_2x + a_3y \quad (22) \\ y &= -a_4 - a_5x + (1-a_6)y \equiv a_4 + a_5x + a_6y \end{aligned}$$

위의 두 식을 x 와 y 에 대하여 나타내면

$$\begin{aligned} x &= b_1 + b_2x + b_3y \\ y &= b_4 + b_5x + b_6y \end{aligned} \quad (23)$$

여기서, $b_1 = (a_3a_4 - a_1a_6)/D$, $b_2 = a_6/D$, $b_3 = -a_3/D$, $b_4 = (a_1a_5 - a_2a_4)/D$, $b_5 = -a_5/D$, $b_6 = a_1/D$, $D = a_2a_6 - a_3a_5$ 이다.

순방향 움직임 파라미터 $\{b_i\}$ 를 위식에 의해 이전 프레임의 $\{a_i\}$ 로부터 계산하고 이 값을 이용하여 이전 영역 분할을 순방향으로 보상하여 현재 분할에 대한 예측 labelling 영상을 얻는다. 순방향 보상 과정은 두 단계로 구성된다. 우선, 움직임이 있는 영역들을 먼저 보상하는데, 이때 두개 이상의 움직임 영역에 의해서 겹쳐 labelling되는 화소는 불확실 영역으로 표시한다. 불확실 영역은 나중의 경계선 결정과정에서 영역의 label값을 결정하게 된다. 그 후, 움직임이 없는 정지 영역의 화소들에 대하여 움직임 영역의 label로 labelling 되어 있지 않는 경우에만 labelling한다.

(2) 불확실 영역의 추출

실제 이전 프레임의 움직임 파라미터들은 추정 오류

를 포함하고 있다. 뿐만 아니라, 이전 프레임의 역방향 움직임 파라미터에서 계산된 순방향 움직임 파라미터는 현재 프레임에서의 실제 순방향 파라미터와는 다르다. 따라서, 이 값들에 의해 보상된 각 영역들의 경계는 실제의 영역 경계와 차이를 나타낸다. 이를 해결하기 위하여 제안 기법에서는 예측된 영역 경계를 중심으로 불확실 영역을 추출하여 이 부분의 정확한 경계값은 경계선 결정 과정에서 결정하도록 하였다. 영역 경계를 중심으로 한 불확실 영역의 설정은 경계부분에서의 영역의 연속성 정도에 바탕을 두는 것이 바람직하지만 본 논문의 실험에서는 단순히 경계 양쪽으로 3화소 두 개의 영역을 할당하였다.

(3) 경계선 결정

불확실 영역에 대한 경계선 결정 과정은 앞절에서

설명된 프레임내 분할에서의 경계선 결정 과정과 동일하다. 다만, 식 (20)에서 영역 평균값 대신 현재의 화소로부터 일정거리 안에 있는 같은 영역의 화소들에 대한 평균값을 사용한다. 이는 프레임내 분할에서의 병합 과정 때문에 하나의 영역이라도 밝기값이 다른 영역이 포함되어 있을 수 있으므로 영역 평균값이 더 이상 그 영역에 대한 특성이 될 수 없기 때문이다.

IV. 실험결과

본 장에서는 제안된 방법으로 동영상을 분할한 결과를 나타내고, 이를 통하여 제안된 방법으로 분할한 결과가 첫째 결합 유사성 척도를 사용함으로써 밝기가 유사하지만 서로 다른 움직임을 갖는 영역들을 정확히



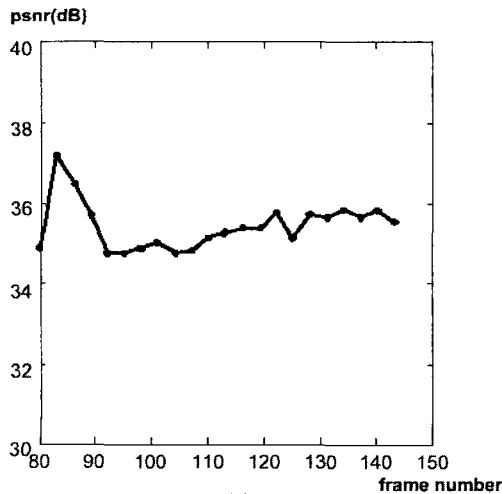
그림 4. "Miss America"에 대한 실험결과: (a) 원영상; (b) 분할결과; (c) 복호화된 영상

Fig. 4. Simulation results for Miss America sequence: (a) original images; (b) segmentation results; (c) reconstructed images.

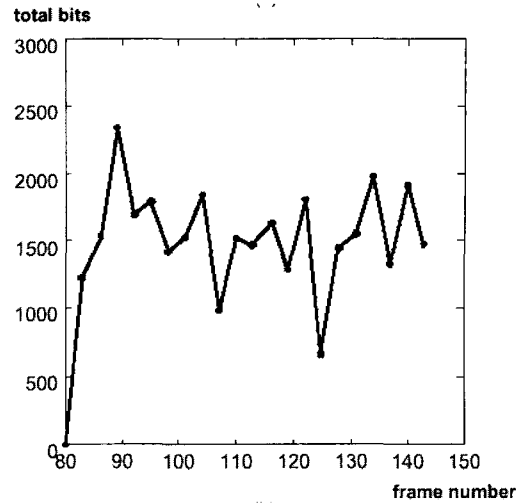
분할할 수 있었으며, 둘째 시간축 연결기법을 기반으로 한 프레임간 분할을 사용하여 연속된 프레임에서 이웃한 영상의 분할결과가 유사성을 나타냄을 보인다. 또한 제안된 동영상 분할 방법을 분할기반 영상 부호화에 적용하여 부호화 효율의 우수성을 보인다.

본 논문의 실험 영상으로는 QCIF 크기의 "Miss America" 및 "Carphone" 시퀀스를 사용하였다. 그리고 영상 단순화시 사용된 structuring element의 크기는 35 x 35로, 밝기 마커의 크기는 250, 움직임 마커의 크기는 70으로 하였다. 영상 부호화시 영상 시퀀스의 프레임율(Frame rate)은 10Hz로 하였으며, "Miss America"는 16kbit/s, "Carphone"은 50kbit/s를 목표 비트율(target bit rate)로 하였다. 움직임 파라미터의 부호화 기법으로는 DPCM을 사용하였다. 그리고 윤곽선 정보량이 전체 부호화의 성능에 미치는 영향을 살펴보기 위해 윤곽선 화소당 필요 bit 수를 가정하면서 실험하였다.

<그림 4>는 "Miss America" 영상에 대한 실험 결과를 보여준다. <그림 4(a)>는 "Miss America" 영상 시퀀스 중 83, 89, 95, 101번째의 원 영상이며, <그림 4(b)>는 제안된 기법을 통해 분할된 결과이다. 사람의 머리부분이 배경영역과 분리되어 분할되었다. 이는 머리와 배경의 경계부분에서 밝기값은 경계가 모호하지만 움직임이 명확히 구분되기 때문에 얻어진 결과로서 결합 유사성 척도의 필요성을 잘 보여준다. 89, 95, 101번째 프레임의 분할 결과가 83번째의 분할과 매우 높은 상관성을 유지함을 알 수 있다. 그러므로, 윤곽선 신



(a)



(b)

그림 5. "Miss America": (a) 신호대 잡음비; (b) 전체 비트수

Fig. 5. "Miss America": (a) PSNR; (b) total bits.

호의 프레임간 부호화 기법을 적용할 경우 매우 높은 부호화 효율을 기대할 수 있다. <그림 4(c)>는 <그림 4(b)>의 분할결과를 이용하여 분할기반 부호화한 뒤, 복호화된 영상이다. 60장의 프레임에 대한 평균 부호화 수치가 <표 1>에 나와있다. 여기서 윤곽선 부호화에 필요한 비트는 윤곽선 화소당 0.7bit를 가정하였다. 프레임당 소요 비트가 1547 비트이며 평균 PSNR이 35.5dB이다. 첫번째 프레임내 부호화된 영상의 PSNR이 34.9dB이었으므로 이는 매우 우수한 화질임을 알 수 있다. 60장의 프레임에 대한 PSNR 및 발생 비트 분포를 <그림 5>에 나타내었다.

표 1. 분할방식에 의한 부호화 결과

Table 1. Performance results by applying proposed segmentation method into coding.

| | PSNR | total bits | motion bits | contour bits | color bits |
|--------------|--------|------------|-------------|--------------|------------|
| Miss america | 35.5dB | 1547 | 168 | 461 | 918 |
| Carphone | 32.5dB | 4782 | 252 | 507 | 4023 |

<그림 6>은 "Carphone" 영상에 대한 실험 결과이다. <그림 6(a)>는 106, 112, 118, 124번째 원 영상이며, <그림 6(b)>의 분할 결과에서 보듯이 전체 영상이 소



그림 6. "Carphone"에 대한 실험결과: (a) 원영상; (b) 분할결과; (c) 복호화된 영상

Fig. 6. Simulation results for "Carphone" sequence: (a) original images; (b) segmentation results; (c) reconstructed images.

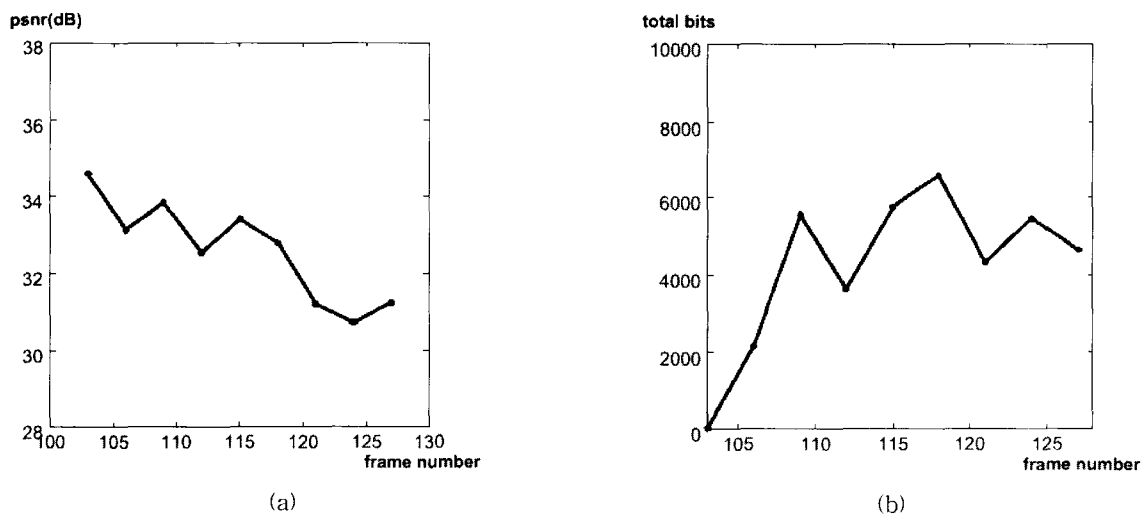


그림 7. "Carphone": (a) 신호대 잡음비; (b) 전체비트수

Fig. 7. "Carphone": (a) PSNR; (b) total bits.

수의 영역으로 분할되었음에도 불구하고 배경영역과 사람 부분이 잘 분할되었으며, 영상간 윤곽선이 높은 상관성을 유지하고 있다. <그림 7>은 "Carphone" 시퀀스 30 프레임에 대한 PSNR 및 발생 비트 분포이며 이의 평균치를 <표 1>에 나타내었다. 프레임당 4782 bits로 평균 32.5dB의 비교적 높은 화질을 나타낸다.

실험결과에서 알 수 있듯이 새로운 결합척도를 이용한 제안 방법이 영역간의 경계를 잘 분할함을 알 수 있고, 또 이전 프레임의 분할된 결과를 현재 프레임의 분할로 연결함으로써 프레임간에 분할된 결과가 매우 높은 상관성을 나타내었다. 이는 제안된 방법을 분할 기반 영상 부호화에 적용할 경우, 윤곽선 정보를 효과적으로 부호화할 수 있음을 의미한다. 분할 기반 영상 부호화에 적용한 실험결과에서 알 수 있듯이 윤곽선 정보를 효과적으로 부호화함으로써 우수한 성능을 나타내었다.

V. 결 론

본 논문에서는 움직임 및 밝기의 결합 척도에 근거한 새로운 영상 분할 기법을 제안하였다. 제안된 기법의 분할 모드는 움직임 벡터장에서 추출된 움직임 마커와 단순화된 영상에서 추출된 밝기 마커 정보로부터 균일한 움직임 및 밝기값을 갖는 결합 마커를 추출하여 이를 seed로 하여 결합 척도에 의해 영역의 경계를 결정하였다. 마지막 단계에서는 추정된 영역의 움직임의 비슷함을 기준으로 영역병합이 수행되었다. 여기서 움직임 마커의 추출을 위해 움직임 벡터장 분할 기법이 새로 제안되었다. 프레임간 분할 모드에서는 영상간의 윤곽선 신호의 상관성을 유지하기 위해 시간축 연결 기법에 의한 분할 기법을 제안하였다. 이전 프레임의 움직임 정보를 이용하여 현재 프레임의 분할을 예측하고, 순방향 움직임 보상의 오류를 수정하기 위한 불확정 영역을 설정한 다음, 이 부분에 결합 척도에 근거한 영역 결정 기법을 적용하여 영상을 분할하였다. 보의 실험 결과를 통하여 제안된 기법이 움직임 보상에 매우 적합하게 영역을 분할하며, 윤곽선 신호의 프레임간 상관성이 매우 높게 유지됨을 살펴보았다.

참 고 문 헌

[1] M. Hotter, "Object-oriented analysis-synthesis

coding based on moving two-dimensional objects," *Signal Processing: Image Comm.*, Vol. 2, pp. 409~428, Dec. 1990.

- [2] J. Ostermann, "Object-oriented analysis-synthesis coding based on the source model of moving flexible 3-D objects," *IEEE Trans. on Image Processing*, Vol. 3, No. 5, pp. 705~711, Sept., 1994.
- [3] P. Gerken, "Object-oriented analysis-synthesis coding of image sequences at very low bit rates," *IEEE Trans. on CSVT*, Vol. 4, No. 3, pp. 228~235, June 1994.
- [4] P. Salembier, M. Pardas, "Hierarchical morphological segmentation for image sequence coding," *IEEE Trans. on Image Processing*, Vol. 3, No. 5, pp. 639~651, Sept. 1994.
- [5] P. Salembier, M. Kunt, "Size-sensitive decomposition of images with rank order sequence coding," *Signal Processing*, Vol. 27, No. 2, pp. 205~241, May 1992.
- [6] Jae Gark Choi, Si-Woong Lee and Seong-Dae Kim, "Spatio-temporal video segmentation using a joint similarity measure," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 7, No. 2, pp. 279~286, April 1997.
- [7] Munchurl Kim, Jae Gark Choi, Daehee Kim, Hyung Lee, Myoung Ho Lee, Chieteuk Ahn and Yo-Sung Ho, "A VOP Generation Tool: Automatic Segmentation of Moving Objects in Image Sequences Based on Spatio-Temporal Information," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 9, No. 8, pp. 1216~1226, Dec. 1999.
- [8] Injae Lee, Won-Sik Cheong, Kyuheon Kim, and Jinwoong Kim, "Precise extraction of moving objects in video sequences for semi-automatic segmentation," *International Workshop on Advanced Image Technology*, pp. 103~107, Hualien, Taiwan, Jan. 16-19 2002.
- [9] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based immersion simulations," *IEEE Trans. on PAMI*, Vol. 13, No. 6, pp. 583~598, June 1991.

저 자 소 개



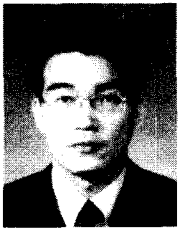
崔在覺(正會員)

1961년 5월 9일생. 1980년 3월~1984년 2월 경북대학교 전자공학과 졸업(공학사). 1985년 3월~1987년 2월 한국과학기술원 전기 및 전자공학과 졸업(공학석사). 1992년 3월~1997년 8월 한국과학기술원 전기 및 전자공학과 졸업(공학박사). 1987년 2월~1998년 2월 한국전자통신연구원 선임연구원. 1998년 3월~2001년 8월 경일대학교 제어계측공학과 조교수. 2001년 9월~현재 동의대학교 컴퓨터공학과 조교수. <주관심분야 : 영상처리, 영상 및 멀티미디어 통신, 워터마킹 등>



南在烈(正會員)

1983년 2월 : 경북대학교 전자공학과 졸업(공학사). 1985년 2월 : 경북대학교 대학원 전자공학과 졸업(공학석사). 1991년 5월 : Univ. of Texas at Arlington 전기공학과 졸업(Ph.D). 1985년 5월~1995년 2월 : 한국전자통신연구소 선임연구원. 1995년 3월~현재 : 계명대학교 컴퓨터공학과 부교수. <주관심분야 : 영상처리, 멀티미디어통신, 멀티미디어 응용>



李時雄(正會員)

1991년 2월 : 경북대학교 전자공학과 졸업(공학사). 1993년 2월 : 한국과학기술원 전기및전자공학과 졸업(공학석사). 1997년 8월 : 한국과학기술원 전기및전자공학과 졸업(공학박사). 1995년 5월~2000년 3월 : 삼성전자반도체. 2000년 4월~현재 : 한밭대학교 조교수. <주관심분야 : 영상처리, 영상압축, 컴퓨터비전>