

시공간 정보를 이용한 영상시퀀스의 영역화

正會員 방 건*, 유 지 상**

Spatio-Temporal Segmentation of Image Sequence

Gun Bang*, Jisang Yoo** *Regular Members*

※이 연구는 정보통신부 대학기초연구지원사업 연구과제의 지원으로 수행된 것임.

요 약

본 논문에서는 시공간 정보를 이용한 영상시퀀스의 영역분할기법과 각 프레임내의 분할된 영역을 다시 전경과 배경 두 부분으로 나누는 움직임 병합기법을 제안하였다. 먼저 영역분할 기법에서는 시공간 정보를 watershed 기법의 입력으로 사용하여 초기 분할영상을 얻었으며 분할된 영역의 수가 많아지는 것을 피하기 위해 그래프이론에 근거한 그래프 병합기법을 적용하였다. 이 때 시공간 정보는 영상의 공간구배(spatial gradient)와 시간구배(temporal gradient)를 적절히 추정함으로써 얻을 수 있었다. 결과로써 얻은 영역들을 움직임이 있는 전경과 움직임이 없는 배경 부분으로 나누기 위해서 움직임 병합기법을 적용하였으며 움직임의 기준은 현재 프레임과 전프레임의 각 영역에 있는 화소들의 평균 절대차이값이 이용되었다. 실험 결과 영상시퀀스의 각 프레임내의 의미 있는 전경에 대해 거의 정확한 경계선을 찾을 수 있었다.

ABSTRACT

An automatic segmentation algorithm which uses both spatial and temporal information and a motion fusion process which divides the contents of each frame of image sequence into two parts, foreground(moving) and background(static) objects are proposed in this paper. The proposed segmentation algorithm is based on the watershed algorithm with the magnitude of spatial-temporal gradient as its input. A region merging process based on graph theory is iteratively applied to the result of the watershed algorithm to avoid over-segmentation. A motion fusion process is also applied to the obtained spatio-temporal segmentation result for dividing an image into moving foreground and static background objects. The average absolute difference of gray levels in each region is used for fusion criterion. Experimental results show that the proposed algorithm gives almost exact contour positions of semantically meaningful foreground objects in test sequences.

*나대기술연구소 연구원
 **광운대학교 전자공학과 교수
 論文番號: 97019-0113
 接受日字: 1997年 1月 13日

I. 서 론

영상영역화는 영상처리분야에서 가장 중요하고 동시에 어려운 기술중의 하나이다. 영역화란 명암도, texture, 움직임 등의 특성에 따라 영상을 그 성질이 유사한 영역으로 나누는 것을 의미한다[1]. 현재까지 많은 영역화 기법들이 제안되었지만 모든 환경에 동시에 좋은 성능을 가지는 기법은 아직 존재하지 않는다. 다시 말하자면 최적의 영역화 기법은 주어진 응용에 따라 또는 주어진 영상의 구조에 따라 달리 설계가 되어야 한다는 것이다.

최근 멀티미디어 및 통신망 기술의 개발에 힘입어 디지털 영상 서비스에 대한 관심이 크게 부각되어 왔다. 특히 이동 통신 및 PSTN 등과 같은 매우 제한된 채널 용량 하에서 영상 서비스에 대한 요구가 크게 증대됨에 따라 초저속 영상부호화에 대한 기술 개발이 절실히 필요하게 되었다. 초저속 영상부호화를 위해서는 기존의 방법과는 다른 새로운 기술들이 개발되어야 하는데 주로 영역기반 부호화 기술들이 많이 거론되고 있다[2][3][4].

영역기반 부호화 방법에서는 화질에 가장 영향을 크게 미치는 블록화 현상이 없고 동시에 객체 지향적인 다양한 처리를 할 수 있다는 장점이 있다. 영역기반 부호화는 영상을 의미 있는 영역으로 분할을 한 후에 그 영역을 기반으로 부호화를 수행하는 방법으로 영역 분할의 성능이 전체 알고리즘의 성능을 좌우하게 된다. 또한 영상 분할 기법은 영상 분석에 가장 기본적인 기술로 초저속 부호화의 전처리과정으로서 사용될 뿐만 아니라 영상 인식, 영상 해석 등의 많은 응용분야에 활용될 수 있는 매우 중요한 기술이다.

본 논문에서는 영상시퀀스에 대해 시공간적 특성을 동시에 고려한 bottom-up 방법의 영상 영역분할 기법을 제안하였다. 주어진 영상시퀀스의 각 프레임을 많은 수의 작은 영역으로 나눈 뒤 서로 상관 관계가 높은 영역들을 병합해 나감으로써 점차적으로 적은 수의 의미있는 영역으로 분할하는 기법이다.

각 프레임에서 의미없는 세부 영역을 가능한 줄이기 위해 각 프레임을 단순화하는 전처리 과정을 먼저 적용하였으며 이때 형태학적(morphological) 연산자가 이용되었다. 단순화된 프레임의 시간구배와 공간구배를 추정하여 여기에 watershed 기법을 적용하여

초기 영역분할 결과를 얻었다. Watershed 기법은 구배영상을 이용하여 영상 내에서 명암도가 가장 낮은 화소부터 시작하여 점차 명암도를 높여 가면서 분수령(crest-line)으로 둘러싸인 부분들을 찾아 그 내부를 하나의 영역으로 분할하는 알고리즘이다[5][6]. 다음 단계로써 초기 분할된 영역의 수를 줄이기 위해 분할된 각 영역간의 명암도 차이와 움직임 차이를 동시에 이용하여 상관도가 높은 영역들을 병합하는 그래프 병합 기법을 적용하였다.

본 논문에서는 또한 분할된 영역들의 움직임을 추정하여 모든 영역들을 움직임이 있는 영역과 움직임이 없는 영역으로 병합하는 움직임 병합기법을 제안하였다. 보통 움직임이 있는 영역이 전경으로 분리되고 움직임이 없는 영역은 배경으로 분리된다. 전경에 보다 큰 의미가 부여되고 따라서 전경과 배경 두 부분으로 영역화되면 부호화시 중요도에 따라 각각의 영역에 서로 다른 비트율을 할당함으로써(scalable transmission coding) 초저속 전송시에 효율을 높일 수 있다는 장점이 있다. 영역기반 부호화 기술을 이용하는 MPEG-4(moving picture experts group)에서도 부호화 기술의 전처리 과정으로 영상시퀀스의 각 프레임을 전경과 배경으로 나누는 과정을 포함하고 있다. 본 논문의 구성을 보면 먼저 II장에서는 영역화 기법에 사용된 여러 가지 형태학적 연산자들을 소개하였으며, III장에서는 시공간 정보를 이용한 영역분할 기법을, IV장에서는 움직임 병합 기법을 각각 설명하였다. 마지막으로 제 V장에서는 실험 및 고찰을 보였다.

II. 형태학적 연산자(Morphological Operators)

2.1 단순화를 위한 형태학적 필터

형태학은 영상의 중요한 형태소인 크기, 모양, 대비, 연결성 등을 다루기 때문에 영상을 분할하는데 매우 유용하게 쓰인다. 기본적인 형태학적 연산자는 불림(dilation)과 녹임(erosion)이며 이치형태소(flat structuring element)를 가지는 불림과 녹임 연산자는 다음과 같이 정의된다[7].

$$\text{불림} : \delta_n(f)(x) = \text{Max}\{f(x-k), k \in M_n\}$$

$$\text{녹임} : \epsilon_n(f)(x) = \text{Min}\{f(x+k), k \in M_n\}$$

여기서 $f(x)$ 는 입력영상이고 x 는 2차원 좌표점을 의미하며 M_n 은 크기가 n 인 이치형태소(Flat structuring element)를 나타낸다. 불림과 녹임을 이용하여 열림(opening)과 닫힘(closing) 연산자도 다음과 같이 정의할 수 있다[7].

$$\text{열림} : \gamma_n(f) = \delta_n(\epsilon_n(f))$$

$$\text{닫힘} : \varphi_n(f) = \epsilon_n(\delta_n(f))$$

대부분의 형태학적 필터는 불림과 녹임의 연산을 이용하여 정의된다. 하지만 기본적인 형태학적 필터는 영상의 밝은 부분이나 어두운 부분을 확장시키거나 축소시킴으로서 입력영상의 물체를 변화시키며 따라서 적용되는 형태소의 크기나 모양에 따라 원영상의 의미 있는 물체의 경계선의 위치가 변할 수 있으며 이것은 영상의 경계선을 유지하면서 단순화시키려는 단순화과정의 목적에 위배된다. 그러므로 물체의 경계는 보존하면서 영상을 단순화시킬 수 있는 연산자를 찾는 것은 매우 중요하다. 이를 위에서 입력 영상 f 에서 유지하고 싶은 물체들이 있으면 이 물체들로 이루어진 참조영상(reference image)을 만들어 다음과 같이 크기 1의 지형학적인 불림(geodesic dilation of size one)과 녹임(geodesic erosion of size one)을 정의할 수 있다[8].

$$\text{크기 1의 지형학적 불림} : \delta^{(1)}(f, r) = \text{Min}\{\delta_1(f), r\} \quad (1)$$

$$\text{크기 1의 지형학적 녹임} : \epsilon^{(1)}(f, r) = -\delta^1(-f, -r) \quad (2)$$

여기서 r 은 참조영상을 나타내며 불림이나 녹임의 결과가 참조 영상에 있는 물체의 경계선을 유지하게 된다[9]. 무한크기의 지형학적 불림과 녹임은 크기 1의 지형학적 불림이나 녹임의 반복적인 형태로 정의되며 개조 불림(reconstruction by dilation)과 개조 녹임(reconstruction by erosion)으로 불린다.

$$\text{개조불림} : \gamma^{(rec)}(f, r) = \delta^{(\infty)}(f, r) = \dots \delta^{(1)}(\dots \delta^{(1)}(f, r) \dots, r)$$

$$\text{개조녹임} : \varphi^{(rec)}(f, r) = \epsilon^{(\infty)}(f, r) = \dots \epsilon^{(1)}(\dots \epsilon^{(1)}(f, r) \dots, r)$$

그러므로 개조열림(opening by reconstruction)과 개조닫힘(closing by reconstruction)은 다음과 같이 정의된다[8].

$$\text{개조열림} : \gamma^{(rec)}(\gamma_n(f), r) \quad (3)$$

$$\text{개조닫힘} : \varphi^{(rec)}(\varphi_n(f), r) \quad (4)$$

본 논문의 단순화 단계에서는 식 (3)의 개조열림 연산자를 적용하고 그 결과에 식 (4)의 개조닫힘 연산자를 적용함으로써 정의되는 개조열림닫힘(open_close by reconstruction) 연산자가 사용되었다. 개조열림 연산자의 참조영상으로는 f 가 사용되고 그 결과는 다시 개조닫힘 연산자의 참조영상으로 사용되었다.

2.2 형태학적 구배 연산자(Morphological Gradient Operator)

Watershed 기법의 입력이 될 영상의 공간구배(spatial gradient)를 추정하기 위하여 다음에 정의되는 형태학적 구배 연산자를 이용하였다[8].

$$\text{형태학적 구배 연산자} : G_s(x) = \delta_1(f(x)) - \epsilon_1(f(x)) \quad (5)$$

III. 시공간 정보를 이용한 영역분할 기법

Bottom-up 방식의 분할 방법은 조그마한 영역의 분할에서부터 점차적으로 그 분할 영역을 크게 하여 분할을 수행하는 방식을 말한다. 본 논문에서 제안한 기법은 단순화, 시공간 구배의 추정, watershed 기법과 그래프 병합 단계 등 크게 4가지 단계로 구성된다. 그림 1에 제안된 기법의 구성도를 보였다.

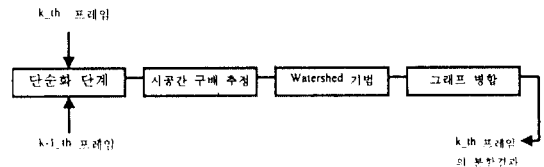


그림 1. 시공간 정보를 이용한 영상분할 기법의 구성도
Fig. 1 Structure of the image segmentation algorithm based on spatio-temporal information

3.1 단순화(simplification) 단계

단순화의 목적은 영상내의 중복도를 높여주며 또한 영상에서 잡음 등을 포함한 불필요한 세부영역을 제거하는 것이다. 정확한 영역분할을 위해서는 단

순화 단계에서 원하는 영역의 경계는 보존하면서 영역을 평활화가 이루어져야 하고 원하지 않는 세부 영역들을 제거할 수 있어야 한다.

먼저 8단계 Lloyd-Max 히스토그램 양자화를 통하여 영상을 평활화 하였다. 여기서 Lloyd-Max 양자화는 영상의 히스토그램 분포에 대해 균일하게 8부분으로 나누어 각 부분의 평균값을 그 부분의 대표값으로 설정하여 양자화 하는 방식이다. 양자화를 통해서 나온 영상은 어느 정도 영상의 잡음이 제거되며 단순화가 되었지만 영상내에 불필요한 요소들은 여전히 남아 있게 된다.

이러한 문제를 해결하기 위해 영역의 경계를 유지하면서 원하지 않는 세부영상을 제거할 수 있는 연산을 작용하여야 하며 이를 위해서 중간값 필터, 스택 필터, 형태학적 필터 등 비선형 필터[9]들이 효과적으로 사용될 수 있다. 본 연구에서는 II장에서 정의된 개조열림단합의 형태학적 필터를 사용하였다.

3.2 시공간 구배 추정

단순화된 영상에 2장의 식 (5)에서 정의된 형태학적 구배 연산자를 적용하여 영상의 공간적 구배를 구한다. 형태학적 구배는 영상에서 명암도의 차이를 제공하며 명암도의 차이가 클수록 구배값도 커진다. 따라서 물체와 배경의 명암도 차이가 거의 없을 경우에는 단순화 단계에서 이미 물체가 배경으로 병합되며 이 경우에는 공간적 구배에 의해서 그 물체의 경계선을 찾을 수 없게 된다. 이 문제를 해결하기 위해서 3D 영상의 시간적 정보인 영역의 움직임을 공간적 구배와 함께 고려하였다. 시간적 정보는 3D 영상의 연속된 두 개의 프레임에 이용하여 시간적 구배를 추정함으로써 얻을 수 있다. 본 연구에서는 다음과 같이 시간적 구배를 정의하였다.

$$\text{시간적 구배: } G_t(x, y) = |f_k(x, y) - f_{k-1}(x, y)| \quad (6)$$

여기서 $f_k(x, y)$ 는 k 번째 프레임의 (x, y) 화소의 명암값이고 $f_{k-1}(x, y)$ 는 $(k-1)$ 번째 프레임의 (x, y) 화소의 명암값이다.

식 (5)와 식 (6)에서 정의된 공간적 구배와 시간적 구배를 적절히 조합하여 다음 단계인 watershed 기법의 입력으로 사용하게 된다. 여기서는 각 화소에서

공간적 구배와 시간적 구배의 최대값을 취하였으며 여기서는 이것을 시공간 구배 G 라고 부르기로 한다.

$$\text{시공간 구배: } G(x, y) = \text{Max}\{G_t(x, y), G_s(x, y)\}$$

여기서 G_t 는 시간적 구배영상이고 G_s 는 형태학적 구배를 이용한 공간적 구배영상을 나타낸다.

3.3 Watershed 기법

시공간 구배영상은 watershed 기법의 입력으로 사용된다. Watershed 기법은 입력영상의 명암도가 최소인 화소에서부터 점차 명암도를 크게 하면서 산등성이로 둘러싸인 부분을 찾아 하나의 영역으로 분할하여 주는 기법이다[8][13][14][15]. 이 때 인접된 두 영역이 만날 때에는 댐을 쌓아주어 이 두 영역을 구분하여 주게 된다. 이 기법은 크게 초기화 단계, 침수 단계로 구분된다. 이를 정리하면 다음과 같다.

3.3.1 초기화 단계

초기 구배 영상은 gray-tone이 0인 지역부터 시작하게 된다. 이러한 지역은 구배 영상전역에 걸쳐 곳곳에 분포해 있으며 이를 보통 최소 지역이라 부른다. 최소지역들이 초기에 레이블이 매겨지면서 초기화가 이루어진다. 또한 gray-tone이 한 레벨씩 높아질 때마다(침수 단계) 이 레벨에서의 최소 지역들이 있는가를 검사하게 된다.

3.3.2 침수 단계

이 단계는 최소지역들을 둘러싸고 있는 점들을 레이블된 최소지역으로 포함시키는 과정이다. 이 때 점들을 어느 영역에 포함해야 하는가는 현재 영역과 이 점들과의 직선 거리를 계산하여 가장 짧은 거리를 가지는 영역으로 결정되게 된다.

각 watershed는 침수가 끝난 후 하나의 영역으로 구분된다. 입력으로 사용되는 구배영상이 매우 복잡하고 watershed 기법이 이에 민감하게 반응함으로 조그마한 영역을 포함하여 매우 많은 영역들이 분할된 결과로 만들어지며 이를 해결하기 위하여 마커 추출을 이용하는 등 여러 방법들이 제시되었지만[8] 본 논문에서는 이 문제의 해결방법으로 그래프 이론에 의한 병합 방법을 제안하였다.

3.4 그래프 병합 기법

그래프 병합 기법에서는 분할된 영역의 수를 줄이면서 더욱 의미 있는 영역들로 영상을 분할하기 위하여 유사한 특성을 가지는 영역들을 병합하게 된다. 이를 위해 영역 분할된 영상은 그래프 형태로 각 영역과 영역의 경계를 정확하게 표현한다.

그래프는 edge와 vertex로 구성된다. 여기서 vertex는 분할된 각 영역을 나타내며 edge는 영역간의 연결성을 나타낸다. Watershed 기법의 결과로써 얻은 분할된 영상을 입력으로 받아 각 영역을 vertex로 표시하여 주고 이 영역들을 edge로 이어주게 된다. 이때 인접 그래프 표시 방법을 이용하면 명확하게 표시할 수 있다[10][13][14][15].

영역의 병합시 병합 요소를 결정하는 일은 매우 중요하며 이에 따라 병합의 결과가 다르게 나타날 수 있다. 본 논문에서는 병합하려는 영역의 시공간 정보를 모두 고려하기 위하여 각 영역에 대해 공간적 정보인 평균 명암도 R_s 와 시간적 정보인 현재 프레임과 과거 프레임 사이의 영역의 평균 절대차이값 R_t 를 적절히 조합하여 병합의 기준으로 사용하였다. 이러한 병합의 기준을 통하여 병합시 이웃 영역과의 유사도(similarity)를 계산하게되며 값이 작을수록 유사도가 높다는 것을 의미한다. 영역 분할된 프레임에서 영역 i 와 이웃 영역 j 간의 유사도를 S 라 하면 다음과 같이 계산될 수 있다.

$$S = |(R_t + \beta R_s)_i - (R_t + \beta R_s)_j| \quad (7)$$

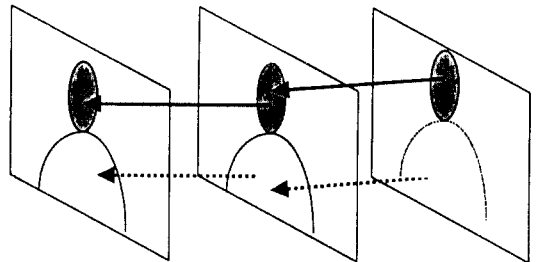
여기서 β 는 R_t 와 R_s 값의 크기를 고려하여 정해지며 R_t 와 R_s 가 유사도 계산시 비슷한 비중을 갖도록 하는 균등화 가중치 상수가 된다. 실험 결과 대부분의 영상에서 β 는 1 보다 작은 양의 실수를 취하게 된다. 만약 β 값을 크게 해주면 병합시 공간적 정보의 영향을 시간적 정보 보다 크게 받게 된다. 반대로 이 값을 훨씬 작게 해주면, 시간적 정보가 병합 과정에서 절대적인 영향을 갖게 된다. 어떤 영역이 이웃 영역과 병합이 되면 병합된 영역에서는 R_t 와 R_s 의 값이 새로이 계산되어진다. 병합은 모든 영역에 대해 수행되고 병합되어야 하는 영역의 크기를 크게 하면서 반복적으로 수행된다.

영역의 병합시 병합의 대상이 되는 영역은 영역의

크기 N_i 가 N_1 의 초기값에서 시작하여 단계적으로 n 만큼씩 증가하면서 정해지고 각 단계에서 N_i 이하의 영역이 영상 내에 존재하면 병합의 대상이 된다. 각 단계에서 전체의 영역에 대한 병합이 끝난 후에는 그래프의 구조는 갱신되어야 한다.

IV. 움직임 병합 기법

영역기반 부호화에서는 영상을 의미 있는 물체로 이루어진 전경과 움직임이 없는 배경으로 분할하는 것이 매우 중요하다. 이는 앞에서 설명한 scalable transmission coding의 근간이 될 수 있기 때문이다. 따라서 3장의 결과로서 얻은 시공간 분할된 영역을 가지고 움직임을 기준으로 움직임이 있는 전경과 움직임이 없는 배경으로 영상을 분할하기 위하여 움직임 병합 기법을 제안하였다.



좌중, k'th 프레임 결과 현재 k'th 프레임 결과 (k-1)th 프레임 결과

그림 2. 과거 프레임 정보를 이용한 움직임 병합

Fig. 2 Motion fusion by using previous frame information

그림 2에서 점선으로 둘러싸인 영역은 각 프레임내에서 움직임이 없는 배경으로 결정된 부분을 나타내고 있다. 이때 k 번째 프레임의 전경을 결정하기 위해서 $k-1$ 번째 프레임의 전경에 관한 정보를 이용한다. 그림에서 현재 k 번째 프레임에서는 머리 부분이 움직임이 없는 배경으로 결정되었기 때문에 전체 움직임 병합의 결과는 매우 좋지 않음을 알 수 있다. 따라서 과거 프레임의 정보를 이용하여 현재 프레임의 결과를 보정하면 그림 2의 최종 k 번째 프레임의 결과와 같이 좋은 결과를 얻을 수 있다.

본 논문에서는 움직임 병합이 두 단계로 이루어지는데 첫 번째 단계에서는 분할된 영상에서 각 영역의

움직임을 추정하여 전경과 배경을 결정하고 두 번째 단계에서는 과거의 정보를 이용하여 첫 번째 단계에서 얻어진 결과를 보강한다.

[1 단계] 각 영역의 움직임을 구한다. 프레임 간의 움직임을 구하는 방법에는 여러 가지가 있으나[11] 여기서는 앞에서 정의한 R_t 를 가지고 움직임을 기준으로 삼았다. 즉 정해진 문턱값보다 큰 R_t 를 가지는 영역들은 움직임을 있는 전경으로, 문턱값보다 작은 R_t 를 갖는 영역들은 모두 움직임을 없는 배경으로 병합되는 것이다.

[2 단계] 과거 프레임의 정보를 이용하여 첫 번째 단계에서 결정된 전경에 대해 보상을 하는 단계로서 다음의 세 가지 가능성을 고려하였다.

- 가. 과거 프레임과 현재 프레임이 모두 전경이거나 배경인 경우:반드시 전경이거나 배경
- 나. 과거 프레임에서는 배경이고 현재 프레임에서는 전경인 경우:배경으로 될 가능성을 조사
- 다. 과거 프레임에서는 전경이고 현재 프레임에서는 배경인 경우:전경으로 될 가능성을 조사

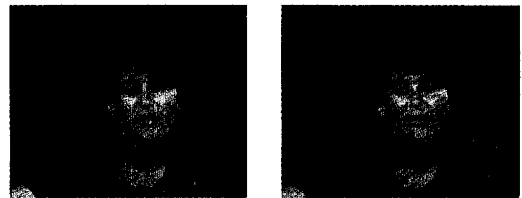
가.의 경우는 확실히 움직임을 있는 전경이거나 움직임을 없는 배경이다. 나.의 경우 이 부분에 해당되는 현재 프레임의 영역들에 대해서만 문턱값을 다시 계산하여 배경으로 결정되면 이 부분만을 배경으로 처리한다. 다.의 경우는 나.의 경우와 동일한 방법으로 문턱값을 구하여 전경과 배경을 구분한다. 단, 과거 프레임에 전경이었던 부분과 현재 프레임에서 배경인 부분이 서로 비슷한 위치와 모양을 가지고 있다면 문턱값과는 무관하게 전경으로 분리한다. 그 이유는 움직임을 없는 배경을 그대로 유지하는 것보다는 전경으로 분리함으로써 현재 프레임에서 전경을 더욱 의미 있게 만들기 때문이다.

V. 실험 및 고찰

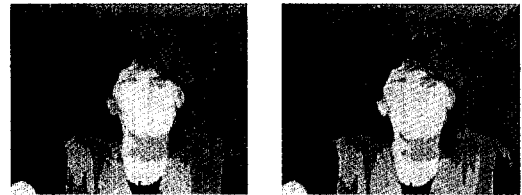
본 논문에서는 그림 3 (가)에 보인 10Hz QCIF 형태의 Miss America 영상 시퀀스에 대해 실험을 했다. 다음은 그 중 103 번째 프레임과 106 번째 프레임에

대한 결과를 보이고자 한다. 실험은 크게 두 가지 구분할 수 있는데 첫 번째는 시공간구배를 이용한 영역 분할이고 두 번째는 분할된 영상에서 움직임을 기준으로 전경과 배경으로 분할하는 것이다.

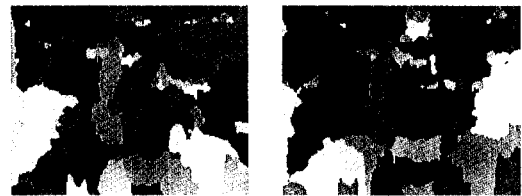
시공간 영상분할에서 단순화를 위하여 8 단계 Lloyd-Max 히스토그램 양자화 기법[1]을 적용한 후 그 결과에 9×9 의 형태소를 가지는 개조열림닫힘 연산자를



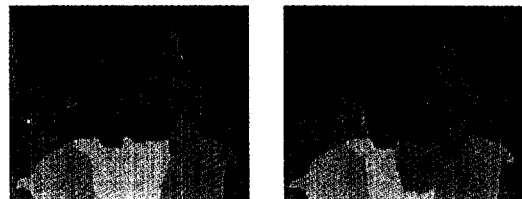
(가) 원 영상



(나) 단순화된 영상



(다) Watershed 기법을 이용한 초기 분할 영상



(라) 그래프 병합된 결과

그림 3. Miss America 영상 시퀀스를 이용한 영역분할 결과. 좌측에 있는 영상들이 103 번째 프레임을 우측에 있는 영상들이 106 번째 프레임을 나타낸다.

Fig. 3 Experimental results for Miss America sequence. Results in the left column were obtained from 103'rd frame and Results in right column from 106'th frame.

적용하였다. 단순화된 각 프레임을 그림 3의 (나)에 보였다. 3×3 크기의 형태학적 구배연산자를 이용하여 단순화된 영상의 공간구배 G_s 를 구하였고 시간구배는 앞에서 정의한 G_t 를 이용하였다. 이와 같이 구한 구배 영상에 watershed 기법을 적용하여 초기 분할된 영상을 얻었으며 이 결과를 그림 3의 (다)에 보였다.

초기 분할된 영상에는 너무 많은 영역이 존재하기 때문에 그래프 이론을 이용하여 영역의 병합을 시도하였다. 그래프 이론에 의해 각 영역들은 vertex와 edge로 이루어진 그래프 형태로 변환될 수 있다. 이 실험에서는 17 개 이하의 영역이 남을 때까지 그래프 병합 기법을 반복하여 적용하였으며 그 결과는 그림 3의 (라)에 보였다. 이때 한 영역과 이웃 영역의 병합 기준인 (7)식에서 β 값은 0.3으로 정해주었다. 만약 β 값을 크게 해주면 병합시 공간적 정보의 영향을 시간적 정보보다 크게 받게 된다. 반대로 이 값을 훨씬 작게 해주면, 시간적 정보가 병합 과정에서 절대적인 영향을 미친다. 이렇게 얻어진 영상의 결과는 어느 의미 있는 영역들로 분할되어진 것을 볼 수 있다.



그림 4. Miss America 영상 시퀀스를 이용한 움직임 병합 결과. 좌측에 있는 영상들이 103 번째 프레임을 우측에 있는 영상들이 106 번째 프레임을 나타낸다.

Fig. 4 Experimental results of motion fusion for Miss America sequence. Results in the left column were obtained from 103rd frame and Results in right column from 106th frame.

분할된 영상을 전경과 배경으로 분할하기 위하여 움직임 병합도 행하였다. 각 영역에 대해 경계값을 제외한 R_i 의 값을 구하고 이 값을 오름차순으로 정렬을 한 후 차이가 5 이상 되는 영역이 있을 경우 두 영역의 R_i 값의 평균을 구하여 문턱값으로 하였다. 만약 차이가 5 보다 작으면 모든 R_i 값의 평균값을 문턱값으로 하였다. 그림 4의 (가)에 103 번째 프레임과 106 번째 프레임의 분할된 전경의 모습을 각각 보였고 그림 4의 (나)에는 각 프레임의 배경을 나타내었다. 움직임이 거의 없는 영역은 모두 배경으로 병합되었으나 시공간 영상분할의 결과가 반영되어 대체로 의미 있는 Miss America의 상반신 대부분이 전경으로 병합되었음을 알 수 있다.

VI. 결 론

본 논문에서는 3D 영상의 시공간 분할기법과 움직임 병합 기법을 제안하였다. 단순화된 영상의 공간구배와 시간구배에 watershed 기법을 적용하였으며 영역의 수를 줄이기 위하여 그래프 이론을 이용한 병합 기법을 제안하였다. 움직임 병합에서는 영상의 움직임을 기준으로 전경과 배경으로 분할하였으며 이때 전 프레임의 정보도 이용하였다. 이는 움직임 객체의 추출이라는 움직임 병합 기법을 제안함으로써 초저속 부호화에 적용하여 더욱 효율적인 부호화를 가능하게 한다는 점에서 본 연구의 의의는 매우 크다고 할 수 있다. 실험결과 움직임이 적은 영상에서도 전경과 배경의 경계선을 비교적 정확하게 찾아내는 것을 알 수 있었다.

참 고 문 헌

1. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison Wesley, 1993.
2. T. Ebrahimi, E. Reusens and W. Li, "New trends in very low bitrate video coding", *Proc. of IEEE*, vol. 83, No. 6, pp. 877-900, June 1995.
3. M. Kunt, A. Ikonomopoulos and M. Kocher, "Second-generation image-coding techniques", *Proc. of IEEE*, vol. 73, pp. 549-574, Apr. 1985.
4. C. Gu and M. Kunt, "Contour simplification and

motion compensation coding," *Signal Processing: Image Communicatton*, vol. 7, no. 4, pp. 279-296, Nov. 1995.

5. L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations.", *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 13, no. 6, pp. 583-598, June 1991.
6. P. Salembier, L. Torres, F. Meyer, C. Gu, "Region-based video coding using mathematical morphology", *Proc. of IEEE*, vol. 83, no. 6, pp. 883-857, June 1995.
7. J. Serra and P. Soille(eds.), *Mathematical Morphology and Its Applications to Image Processing*, pp. 69-76, Kluwer Academic Publishers. Printed in the Netherlands, 1994.
8. P. Salembier, "Morphological multiscale segmentation for image coding," *Signal Processing*, vol. 38, no. 3, *Special Issue on Nonlinear Signal Processing*, pp. 359-386, Sept. 1990.
9. C. Lantuejoul and F. Maisonneuve, "Geodesic methods in image analysis", *Pattern Reconignition*, vol. 17, 1984, pp. 117-187.
10. I. Pitas and A. N. Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications*, Kluwer Academic Publishers, 1990.
11. D. H. Ballard and C. M. Brown, *Computer Vision*, Prentice Hall, 1982.
12. N. M. Tekalp, *Digital Video Processing*, Prentice Hall, 1995.
13. Young-Kwon Lim, Myoung Ho Lee, Gun Bang, Jisang Yoo: Automatic segmentation based on spatio-temporal information (ETRI description of CE N2); MPEG96/991, Jul., 1996.
14. Jae Gark Choi, Young-Kwon Lim, Myoung Ho Lee, Gun Bang, Jisang Yoo; Automatic segmentation based on spatio-temporal information (ETRI description of CE N2); MPEG97/1284, Sept., 1996.
15. Jae Gark Choi, Young-Kwon Lim, Myoung Ho Lee, Gun Bang, Jisang Yoo; Automatic segmentation based on spatio-temporal information; MPEG97/1369, Nov., 1996.



劉 智 相(Jisang Yoo) 정희원
 1985년 2월: 서울대학교 전자공학과 졸업(공학사)
 1987년 2월: 서울대학교 대학원 전자공학과 졸업(공학석사)
 1993년 5월: Purdue 대학교 전자공학과 졸업(Ph.D.)
 1993년 5월~1993년 9월: Purdue 대학교 전기공학과 연구원
 1993년 9월~1994년 8월: 현대전자산업(주) 산전연구소 선임연구원
 1994년 9월~1997년 8월: 한림대학교 전자공학과 조교수
 1997년 9월~현재: 광운대학교 전자공학과 조교수
 ※주관심분야: 비선형 신호처리, 영상처리, 컴퓨터 그래픽스



方 健(Gun Bang) 정희원
 1995년 2월: 한림대학교 컴퓨터공학과 졸업(공학사)
 1997년 2월: 한림대학교 대학원 컴퓨터공학과 졸업(공학석사)
 1997년 3월: 고려대학교 대학원 컴퓨터학과 박사과정 입학
 1997년 8월~현재: (주)나다가기술연구소 연구원
 1997년 10월: 한국정보과학회 추계학술대회 우수논문 지정, MPEG-Korea 전문위원
 ※주관심분야: 화상 회의, MPEG-4, MPEG-7, 영상 인식, 객체 움직임 추적