

결합 유사성 척도를 이용한 시공간 영상 분할

正會員 최 재 각*, 이 시 응*, 조 순 제*, 김 성 대*

Spatio-Temporal Video Segmentation Using a Joint Similarity Measure

Jae Gark Choi*, Si-Woong Lee*, Soon-Jae Cho*, Seong-Dae Kim* *Regular Members*

요 약

본 논문은 새로운 시공간 영상 분할(spatio-temporal segmentation) 알고리즘을 제안한다. 제안 알고리즘은 morphological filter 및 watershed 알고리즘을 기본 도구(tool)로 사용하고 밝기 정보와 움직임 정보를 동시에 고려하여 영상을 분할한다.

전체적인 분할 단계는 다음처럼 3 단계로 이루어 진다: 결합 마커(joint marker) 추출, 영역 경계 결정(boundary decision) 그리고 움직임 기준 영역 병합(motion-based region fusion). 첫단계로 결합 마커 추출을 행하게 되는데, 결합 마커란 밝기와 움직임이 동시에 균일한 영역으로 영역 분할을 위한 기준 영역(seed region)으로 사용된다. 여기서 간단하고도 효율적인 결합 마커 추출 방법이 제안된다. 둘째 단계로 watershed 알고리즘에 의해 시공간 영역의 경계(spatio-temporal boundary)가 결정된다. 이러한 영역 경계 결정을 위해 밝기 정보와 움직임 정보를 동시에 결합한 새로운 유사성 척도(joint similarity measure)를 제안한다. 마지막으로 움직임 기준 영역 병합을 사용하여 앞단계에서 나누어진 영역들중 밝기가 다르지만 움직임이 동일한 영역을 병합하게 된다. 실험 결과, 시공간 정보를 동시에 사용하는 제안 방식이 시각적으로 의미있는 분할 결과를 나타내었다.

ABSTRACT

This paper presents a new morphological spatio-temporal segmentation algorithm. The algorithm incorporates luminance and motion information simultaneously, and uses morphological tools such as morphological filters and watershed algorithm. The procedure toward complete segmentation consists of three steps: joint marker extraction, boundary decision, and motion-based region fusion. First, the joint marker extraction identifies the presence of homogeneous regions in both motion and luminance, where a simple joint marker extraction technique is proposed.

*한국과학기술원 전기 및 전자공학과
論文番號:96133-0430
接受日字:1996年 4月 30日

Second, the spatio-temporal boundaries are decided by the watershed algorithm. For this purpose, a new joint similarity measure is proposed. Finally, an elimination of redundant regions is done using motion-based region fusion. By incorporating spatial and temporal information simultaneously, we can obtain visually meaningful segmentation results. Simulation results demonstrates the efficiency of the proposed method.

I. 서 론

최근 멀티미디어(multimedia) 및 통신망 기술의 개발에 힘입어 디지털 영상 서비스에 대한 관심이 크게 부각되어 왔다. 특히 이동 통신(mobile communication) 및 PSTN등과 같은 매우 제한된 채널 용량하에서 영상 서비스에 대한 요구가 크게 증대되어 왔다. 따라서 64kbps 이하의 초저전송률 영상 부호화(very low bit-rate video coding)에 대한 기술 개발이 필요하게 되었다[1]. 그러나 H.261이나 MPEG-1 및 MPEG-2등과 같은 기존의 동영상 부호화 기법들은 이동 물체의 모양 및 전역 움직임(global motion)과 무관하게, 영상을 블록으로 나누고 블록단위의 움직임 추정 및 보상을 행하게 된다. 따라서 낮은 전송률에서, 블럭화 현상(block effect), 모서리 떨림 현상(edge buyness), 반점현상(mosquito effect)등과 같은 화질 저하가 재생 영상에 나타나게 된다. 저전송률 영상 부호화의 이러한 문제점을 해결할 수 있는 가능한 방안은 Kunt[2] 등이 제안한 것과 같은 2세대 부호화(the second generation coding)이다. 이러한 방법들은 영상 신호의 통계적 특성만을 주로 고려하는 기존 방법과는 다르게, 인간의 시각 특성에 바탕을 두고 주관적 화질을 중요시하며 아울러 전송 데이터량의 대폭적인 감축을 도모하는 부호화 기법이다.

이러한 부호화 기법들 중 영역기반 부호화(region-based coding) 또는 물체별 부호화(object-based coding) 방법[3][4]이 현재 중요한 차세대 부호화 기법 중 하나로 연구되고 있다. 영역기반 또는 물체별 부호화와 같은 분할기반 부호화는 영상 내용을 주어진 모형(예, 2D/3D 이동물체 또는 밝기가 균일한 영역)에 따라 분할한 뒤, 분할된 영역의 파라미터(예, color, shape and motion)를 부호화하여 전송한다. 따라서 분할된 영역별로 부호화가 행하여 지기 때문에 특히 영역의 경계 부근에서 우수한 재생 화질을 나타낼 수 있다. 영상 분할은 이러한 분할기반 부호화에서 꼭 필요한

도구이며 본 논문의 연구 내용이다.

영상 분할을 위한 기존 방법은 분할 방법에 따라 split and merge 방법[5]에서 부터 morphological 분할[6][7]까지 다양한 방법이 존재하고 영상 분할시 사용되는 분할기준(homogeneity criterion)으로는 움직임(motion), 밝기값(intensity) 및 색차신호(Cb and Cr) 등이 있다. 움직임을 기준으로 한 영상 분할은 움직임 물체의 운동모형을 가정한 뒤 가정된 움직임 모형에 따라 움직임 벡터장을 분할하는 방법이다. 움직임이 영상 분할의 기준이 될 때 다음과 같은 특징이 있다. 분할기반 부호화에서 영역별로 움직임 추정 및 보상을 행한다는 관점에서 본다면 움직임에 따라 영상을 분할하는 것이 보다 적합한 영상 분할 방법이다. 하지만 잘 알려진 것처럼 기존의 움직임 추정 방법들은 물체의 경계부근에서 부정확한 움직임을 추정한다[8]. 따라서 분할 결과가 실제 물체 경계와 일치하지 않게 되고 영상 부호화시 부호화 효율이 떨어지게 된다. 움직임에 따라 영상을 분할한 예로 Hotter의 계층적 분할[9], 운동 모형에 대한 제한(constraint)을 완화함으로써 단계적으로 움직임 벡터장을 분할하는 다단계 분할법[10] 그리고 Hough 변환을 이용하는 분할법[11]등이 있다. 밝기값을 기준으로 영상을 분할하는 것은 움직임 물체의 경계는 밝기값의 경계(intensity boundary)와 일치한다는 가정을 전제로 한다. 따라서 분할 기준이 밝기값이 될 경우 물체 경계가 인간 시각 특성 상의 영상 경계와 일치하게 된다. 하지만 영상이 수 많은 균일한 밝기값의 영역으로 나누어 지게되어 움직임 관점에서 비슷한 움직임을 갖는 영역이 존재하게 된다. 따라서 영역의 모양 및 움직임 정보를 부호화하는데 많은 정보량이 필요하게 된다. 밝기값에 따라 영상을 분할한 예는 Salembier의 morphological filter를 사용한 계층적 분할[6][7]등이 있다.

최근 움직임 기준 및 밝기값 기준 분할 방법의 문제점들을 해결하기 위해 움직임 및 밝기값을 함께 분

할 기준으로 하는 시공간 영상 분할법이 제안되었다. 참고문헌 [12] 및 [13]에서는 영상을 밝기값으로 분할한 뒤(spatial segmentation), 이들 영역을 움직임 기준으로 다시 병합하는(temporal segmentation) 영상 분할법을 사용하였다. 한편 움직임 벡터장과 밝기값을 함께 고려하여 morphological tool로 분할하는 방법이 참고문헌[14] 및 [15]에 소개되었다. 이러한 영상 분할 방법은 밝기값과 움직임 정보를 함께 이용하는 관점에서 움직임이나 밝기값 하나만을 이용하는 영상 분할법보다 우수한 결과를 나타내지만 밝기값 또는 움직임을 이용하여 각각 독립적으로 영상을 분할한 뒤, 이 결과를 조합하는 방식으로 영상분할이 이루어 진다. 따라서 밝기값이 비슷한 물체가 하나의 물체로 병합되는 경우 이를 수정하기가 어렵다.

보다 정확한 영상 분할을 위해서는 밝기값과 움직임 정보를 동시에 고려하는 분할 방법에 대한 연구가 이루어져야 한다. 시공간 정보를 동시에 이용하는 영상 분할(joint spatio-temporal segmentation)이 본 연구 목표이며, morphological filter 및 watershed 알고리즘을 기본 도구(tool)로 사용하고 밝기 정보와 움직임 정보를 동시에 고려하여 영상을 분할한다. 전체적인 분할 단계는 다음처럼 3 단계로 이루어 진다. 첫 단계로 joint marker 추출을 행하게 되는데 joint marker란 밝기값과 움직임이 동시에 균일한 영역으로 영역 분할을 위한 기준 영역(seed region)으로 사용된다. 여기서 간단하고도 효율적인 joint marker 추출 방법이 제안된다. 둘째 단계로 watershed 알고리즘에 의해 시공간 영역의 경계가 결정된다. 이러한 영역 경계 결정을 위해 밝기 정보와 움직임 정보를 동시에 결합한 새로운 결합 유사성 척도(joint similarity measure)를 제안한다. 마지막으로 움직임 기준 영역 병합을 사용하여 앞단계에서 나누어진 영역들중 밝기가 다르지만 움직임이 동일한 영역을 병합하게 된다. 실험 결과, 시공간 정보를 동시에 사용하는 제안 방식이 시각적으로 의미있는 분할 결과를 나타낸다.

본 논문의 구성은 다음과 같다. 2절에서 시공간 영상 분할을 위해 제안된 시공간 결합 유사성 척도에 대해 기술한다. 3절에서 제안된 시공간 영상분할 알고리즘에 대해 설명하고 4절에서 제안된 시공간 영상 분할 방식의 성능 평가를 위한 모의실험 결과를 제시한다. 마지막으로 5절에서 결론을 맺는다.

II. 시공간 결합 유사성 척도

본절에서는 morphological 시공간 영상 분할을 위한 새로운 결합 유사성 척도에 대해 기술한다. 먼저 기존의 결합 유사성 척도[16]의 문제점들을 알아보고 이러한 문제점을 해결할 수 있는 새로운 유사성 척도를 제안한다.

2.1 시공간 결합 유사성 척도의 특성

밝기와 움직임을 동시에 고려한 영상 분할법에서 사용할 수 있는 유사성 척도(joint similarity measure)는 식 (1)처럼 밝기값의 차와 움직임 차의 합으로 표현될 수 있다.

$$\text{similarity} = \alpha \text{ intensity difference} + (1 - \alpha) k \text{ motion difference} \quad (1)$$

여기서 유사성 척도란 영역 R과 영역 R로 병합하려는 화소(x, y) 간의 유사성 정도를 말한다. 그리고 α 는 가중치(weighting factor)로 밝기값의 차와 움직임의 차 사이의 중요도를 조절하는 인자이다. 또 k는 밝기값 차와 움직임 벡터차는 서로 다른 단위를 가지기 때문에 이들 사이에 크기(scaling) 조절을 위한 값(scale factor)이다.

밝기값의 차는 영역 R의 평균 밝기값과 고려중인 화소(x, y)의 밝기값과의 차를 의미하며 다음처럼 표현된다.

$$\text{intensity difference} = |\text{intensity avg. of } R - I(x, y)| \quad (2)$$

여기서 $I(x, y)$ 는 화소(x, y)에서 밝기값이다. 그리고 움직임의 차는 화소(x, y)의 추정된 움직임 벡터와 영역 R의 움직임 파라미터로부터 계산된 화소(x, y)의 움직임 벡터와의 차를 의미하며 다음처럼 표현된다.

$$\text{motion difference} = \delta(x, y) = \sqrt{(d_x(x, y) - d_x^0(x, y))^2 + (d_y(x, y) - d_y^0(x, y))^2} \quad (3)$$

여기서 $(d_x(x, y), d_y(x, y))$ 는 화소(x, y)에서의 추정된 움직임 벡터를 나타내며 $(d_x^0(x, y), d_y^0(x, y))$ 는 영

역 R 의 움직임 파라미터 θ 에 의해 화소 (x, y) 에서 계산된 움직임 벡터를 나타낸다. 움직임 모형을 affine 모형으로 가정할 경우 $(d_x^o(x, y), d_y^o(x, y))$ 는 식 (4)로 표현된다.

$$d_x^o = a_1 + a_2x + a_3y, \quad d_y^o = a_4 + a_5x + a_6y \quad (4)$$

식 (1)과 같은 시공간 결합 유사성 척도를 이용하는 경우 다음과 같은 문제점이 있다.

- 밝기값의 차와 움직임의 차는 서로 다른 단위를 가지기 때문에 단순히 밝기값의 차와 움직임의 차의 가중합(weighted sum)으로 유사성 척도를 정의할 수 없다.
- 존재하는 움직임 추정법[19][20][21]으로 추정된 움직임 벡터는 잘 알려진 것처럼 잡음을 포함하고 있으며 특히 움직임 경계나 변화가 없는 배경(static background)에서 부정확하다. 따라서 영역 경계를 결정할 때, 식 (3)처럼 화소 (x, y) 의 추정된 움직임 벡터를 직접 사용하면 잘못된 경계를 추정할 수 있다.

2.2 제안한 시공간 결합 유사성 척도

앞절에서 추정된 움직임 벡터를 직접 사용하는 결합 유사성 척도의 문제점을 분석하였다. 본 절에서는 이러한 문제점을 극복할 수 있는 새로운 유사성 척도가 제안된다.

먼저 밝기값의 유사성 척도(intensity similarity)로서 식 (2)처럼 단순히 영역 R 의 평균 밝기값을 사용하는 것이 아니라, 보다 정확한 표현을 위해 영역 R 의 다항 근사치(polynominal approximation)을 이용한다. 따라서 밝기값의 유사성 척도는 다음처럼 표현된다.

$$S_R^i(x, y) = |I_{poly}(x, y; \phi) - I(x, y)| \quad (5)$$

여기서 $I_{poly}(x, y; \phi) = p_1 + p_2x + p_3y$ 이며 화소 (x, y) 에서 영역 R 의 1차 다항 근사화된 값이다. 그리고 움직임의 유사성 척도(motion similarity)로서 움직임의 차를 이용하는 것이 아니라 간접적인 움직임 정보로 움직임 보상 오차(displaced frame difference)를 이용한다. 따라서 고려중인 화소 (x, y) 와 영역 R 사이의 움직임의 유사성은 식 (3)과는 달리 다음처럼 정의된다.

$$S_R^m(x, y) = dfd(x, y; \theta) = |I_k(x, y) - I_{k-1}(x - d_x^o(x, y), y - d_y^o(x, y))| \quad (6)$$

여기서 화소 (x, y) 의 움직임 보상 오차를 계산할 때 화소 (x, y) 의 추정된 움직임 벡터 $(d_x^o(x, y), d_y^o(x, y))$ 를 직접 이용하는 것이 아니라 영역 R 의 affine 움직임 파라미터로 화소 (x, y) 의 움직임 벡터를 계산한 뒤, 이를 이용하여 움직임 보상 오차를 계산한다. 따라서 제안된 결합 유사성 척도를 이용하므로써 다음과 같은 장점들이 있다.

- 움직임의 유사성 척도(motion similarity measure)로 움직임 차를 이용하는 것이 아니라 간접적인 척도로 움직임 보상 오차를 이용함으로써 밝기값의 유사성 척도(intensity similarity)와 같은 단위(밝기값의 차)를 갖는다.
- 화소 (x, y) 의 추정된 움직임 벡터를 직접 이용하지 않고 영역 R 의 움직임 파라미터로 부터 계산된 움직임 벡터를 이용하여 움직임 보상 오차를 계산하기 때문에 움직임 경계나 변화가 없는 배경 영역(static background)에서 발생하는 움직임 추정의 부정확성에 기인된 문제를 해결할 수 있다.
- 움직임 보상 오차를 최소화하도록 되도록 영상분할이 이루어지기 때문에 움직임 보상시 보상 효율이 극대화된다.

따라서 시공간 정보를 동시에 이용하는 제안 결합 유사성 척도는 다음처럼 정의된다.

$$S_R(x, y) = \alpha S_R^i(x, y) + (1 - \alpha)S_R^m(x, y) \quad (7)$$

III. 제안된 시공간 영상 분할 알고리즘

3.1 개요

본 절에서는 움직임 정보와 밝기 정보를 동시에 이용하는 시공간 영상분할 알고리즘에 대해 설명한다. 제안된 영상 분할 알고리즘은 morphological filters와 watershed 알고리즘을 기본 도구로 하고 다음처럼 3 단계로 구성된다.

- 시공간 결합 marker 추출(joint marker extraction)

- 경계 결정(boundary decision)
- 움직임 기준 영역 병합(motion-based region fusion)

그림 1에 제안된 시공간 영상 분할 알고리즘(joint spatio-temporal segmentation)의 구성도를 나타내었다. 맨 첫 단계가 joint marker 추출 단계이며, joint marker란 움직임과 밝기가 동시에 균일한 기준 영역(seed region)이다. joint marker 추출결과, 분할할 영역의 갯수와 분할을 위한 기준 영역이 추출된다. 둘째 단계로서 watershed 알고리즘에 의해 시공간 영역의 경계가 결정된다. 마지막으로 부호화 관점에서 중복성이 있는 영역의 경계를 제거하기 위하여 움직임 기준 영역 병합이 이루어진다.

지금부터 본 논문에서는 영역 또는 물체의 움직임을 기술하기 위해 식 (4)에 정의된 affine 모형을 사용한다.

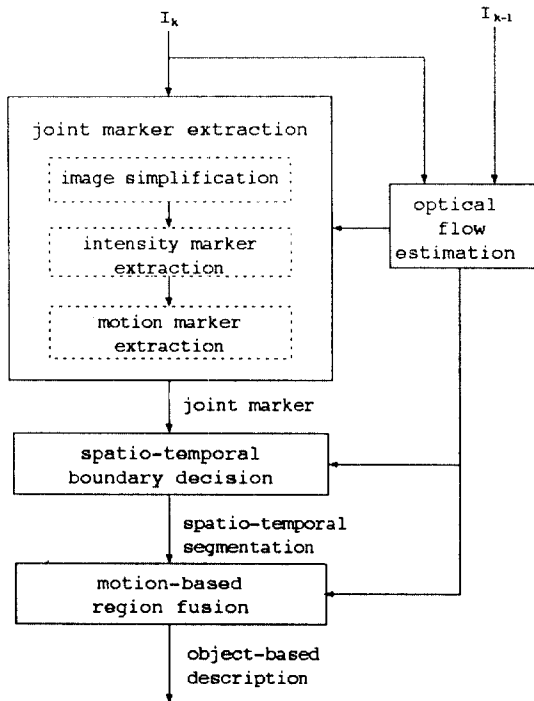


그림 1. 제안된 시공간 영상분할 알고리즘의 구조

3.2 시공간 결합 marker 추출

성공적인 morphological segmentation을 위해서는

시각적으로 중요한 marker의 추출이 매우 중요하다 [14]. 참고문헌 [6]에서 morphological filter로 밝기 영상(luminance image)을 단순화한뒤 크기(size) 및 대조(contrast)를 이용하여 marker를 추출하였다. 이러한 marker 추출은 밝기값이 비슷한 경우, 이웃영역들이 하나의 marker로 추출되고 따라서 잘못된 영역 경계를 분할할 수 있다.

밝기값과 움직임 정보를 동시에 고려하여 marker를 추출한다면 더욱 시각적으로 의미있는 marker 추출이 가능할 것이다. 밝기값과 움직임을 동시에 고려한 결합 marker 추출에 대한 연구가 참고문헌 [14]에서 연구되었지만 여기서는 움직임이 있고 없고의 단순한 구분만을 이용하였다. 따라서 여러개의 이동물체가 존재하는 경우 사람이 직접 marker를 추출하는 것이(manual marker extraction) 불가피하다. 본 논문에서는 밝기와 움직임이 동시에 균일한 결합 marker를 추출하는 간단하고도 효과적인 방법을 제안한다. 그림 1에 나타난 것처럼 결합 marker를 추출하는 과정은 먼저 밝기값이 균일한 intensity marker를 추출한 뒤, intensity marker 내에 존재하는 움직임이 균일한 영역들, 즉 motion marker들을 추출함으로써 시공간 결합 marker들을 추출한다. 그 이유는 움직임은 벡터장이기 때문에 벡터장에서 균일한 영역을 추출하기가 어렵다. 따라서 밝기값이 균일한 영역을 먼저 추출한뒤 이로부터 motion marker를 추출하여 이것을 결합 marker로 하는 것이 훨씬 간단하고 효과적이기 때문이다.

intensity marker를 추출하기 위해 먼저 영상을 단순화한다. 이때 open-close by reconstruction filter를 이용하였다[6]. 이 필터는 주어진 크기보다 작은 영역을 제거하고 나머지 영역의 윤곽선은 보존하는 성질을 갖는다. 단순화된 영상으로부터 밝기값이 균일한 영역(flat region)을 라벨링(labeling)함으로써 intensity marker를 추출하는데 주어진 문턱값보다 큰 영역만을 intensity marker로 추출한다. 그리고 이때 라벨링 방법으로 blob coloring[17]을 사용하였다.

intensity marker 내의 움직임이 균일한 motion marker를 추출할때 다음과 같은 여러가지 상황이 있을 수 있다.

- 한 intensity marker 내에 motion marker가 없는 경우

- 한 intensity marker 내에 하나의 motion marker만 있는 경우
- 한 intensity marker 내에 여러 개의 motion marker 가 있는 경우

따라서 intensity marker내에서 움직임이 균일한 motion marker를 추출할 때 intensity marker 내에 균 일한 움직임을 갖는 영역이 여러개 존재할 수 있기 때문에 이들을 추출하는 다단계 과정(multipass marker extraction algorithm)을 그림 2에 표시하였다. 이 방법은 반복적인 과정으로서 매 반복시마다 이미 생성된 motion marker에 포함되지 않는 화소들 중에서 연결될 수 있는 가장 큰 motion marker를 찾아내는 기법이다. 각 반복 과정에서 수행되는 세부과정은 다음과 같다.

1. 아직 motion marker로 분리되지 않은 화소들중에서 연결될 수 있는 가장 큰 영역 S 를 찾는다. 이 영역의 크기가 임계값 h 보다 작으면 새로운 motion marker가 없는 것으로 간주하고, 새로운 intensity marker에 대해 motion marker를 추출하는 과정을 반복한다. 그리고 한 intensity marker 내에서 이미 생성된 motion marker의 갯수가 정해진 수(일반적으로 3)를 넘는 경우에도, 다단계 방식을 중지한다. 이것은 최대 motion marker 갯수를 제한함으로써, 반복이 지나치게 되풀이되는 것을 방지하기 위함이다.
2. 영역 S 에 대해 이동 벡터장(dense motion field)을 이용하여 affine motion parameter를 구한다. affine 변환을 구하는 과정은 3.4절에서 상세히 설명된다.
3. 앞 단계에서 구한 affine 변환계수를 가지고 S 내의 각 움직임 벡터에 대해 식 (3)에 정의된 δ 를 계산한다. 만약 δ 가 임계값 ϵ 이하이면 그 화소를 새로운 집합 S' 에 포함시킨다. 만약 S' 의 크기가 h 를 넘지 않으면 집합 S' 내부는 더 이상 motion marker로 묶일 수 없으므로 S' 을 반복과정에서 제외시킨 후 새로운 반복과정을 시작한다.
4. S' 내부에서 가장 큰 연결 영역 S'' 을 찾는다. 영역 S'' 의 크기가 h 를 넘지 않으면 이 영역을 반복과정에서 제외시킨다.
5. 만약 영역 S'' 의 크기가 h 보다 크면 S'' 을 motion marker로 등록한 후 새로운 반복과정을 시작한다.

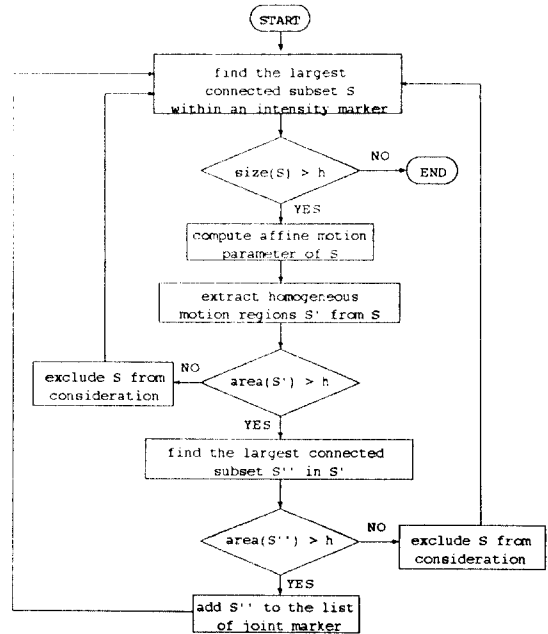


그림 2. 다단계(multipass) 움직임 마커 추출 알고리즘

3.3 시공간 영역 경계 결정

본 절은 앞에서 언급한 joint marker로 부터 정확한 영역의 경계를 결정하는 과정이다. 이때 사용되는 도구는 watershed 알고리즘이며 앞절에서 제안된 시공간 결합 유사성 척도를 사용한다.

Watershed 알고리즘은 주로 지형학 분야에서 연구된 것으로, 영상 화소의 밝기값의 크기를 고도(altitude)로 생각함으로써 영상처리에 응용되었다. 이는 답수지역을 구분하는 분수령을 찾음으로써, 각각의 국부 최소치에 연관된 영역을 분할해 내는 알고리즘이다. 이제까지 분수령을 계산하기 위한 다양한 알고리즘이 제안되었으나, 가장 효과적인 것은 immersion simulation 기법[18]으로서, 이것은 계층적 queue에 의하여 구현된다.

Immersion simulation은 이중 순서의 특성을 가진다. 즉, 유사성이 큰 화소가 유사성이 낮은 화소보다 먼저 병합되며, 또 유사성이 같은 때는 먼저 병합된쪽에 가까운 위치가 먼 위치보다 먼저 병합된다. 이 성질을 이용하여 immersion simulation을 계층적 queue를 사용하여 구현할 수 있다. 계층적 queue는 우선권(priority)이 다른 queue들의 집합이다. 이 각각의 queue

들은 FIFO(first-in-first-out)의 데이터 구조로 되어 있으며, queue에 의해서 처리되는 원소들은 화소의 위치이다. 이 계층적 queue도 이중 순서의 특징을 가진다. 즉, 우선권이 높은 queue에 있는 원소들이 먼저 처리되며, 같은 queue내에서는 먼저 들어간 원소가 먼저 처리된다.

계층적 queue가 immersion simulation을 구현하는 과정은 queue 초기화 과정과 범람(flooding)과정의 두 단계로 구성된다. 초기화는 모든 국부 최소치의 좌표를 queue에 입력하는 과정이다. 즉 marker의 이웃한 화소를 queue에 입력하는 과정이다. 이때, 그들의 유사성 척도의 역수로써 우선권을 계산한다. 그러므로 우선권에 따라 hierarchical queue 중 하나의 queue에 marker에 이웃한 화소의 위치를 입력하게 된다. 범람 과정은 queue에서 화소들을 꺼내는 것으로 구성된다. 만약 그 화소가 아직 marker에 속해 있지 않다면 그 화소의 주위에는 적어도 하나 이상의 이웃한 marker가 존재한다. 그러므로, 이웃한 모든 marker 영역을 검사하여 시공간 결합 척도가 가장 가까운 기준영역(marker)으로 할당한다. 그런 후, 이 화소에 접해있는 화소들 중에서 아직 기준영역에 할당되어 있지 않는 화소들을 queue에 입력한다. 즉, queue에 입력되는 화소들은 주위의 기준영역의 화소를 접하고 있는 화소들이다. 이때 역시 유사성 척도를 계산한 후, 유사성 척도에 따라 hierarchical queue 중 하나에 입력한다. 이 범람 과정을 모든 화소들이 영역에 할당될 때까지 계속한다.

일단 화소가 어떤 영역으로 할당되면 새로운 화소에 대해 정확한 유사성 척도를 계산하기 위하여 그 영역의 움직임 모형과 밝기에 대한 다항 계수 모형을 새롭게 계산해야한다. 그러나, 영역 확장시 매 화소가 영역으로 할당될 때마다 영역의 다항 계수 및 affine 변환을 새로이 계산한다면 너무 계산량이 많아지기 때문에 본 논문에서는 n 개의 화소가 영역에 새로이 추가될때 마다 다항 계수 및 affine 변환을 다시 계산하였다. 몇 개의 화소가 영역에 새로이 추가된다 하더라도 그 영역의 움직임 모형과 밝기값 모형이 갑자기 변화되지는 않기 때문에 이러한 모형 갱신(update)이 가능하다. 영역의 밝기값의 다항 계수 모형(식 (5))은 영역의 밝기값과 근사화된 다항 계수값 사이의 자승 오차 합을 최소화하도록 계산하였다. 또한 영역의

affine 움직임 파라미터는 영역내의 이동 벡터들로부터 계산하였는데 식 (8)에 정의된 오차함수를 최소화하도록 하였다. 그리고 추가되는 화소의 움직임 벡터와 영역의 움직임 파라미터로부터 계산된 움직임 벡터와의 차이가 임계치보다 클 경우에는 이 화소의 움직임 벡터는 신뢰성이 적기 때문에 affine 변환 update 시 제외하였다.

3.4 움직임 기준 영역 병합

앞절에서의 영상 분할은 시공간 결합 유사성 척도를 사용하여 움직임과 밝기가 동시에 균일한 영역으로 영상을 분할한다. 따라서 밝기에 대한 다항계수 (polynomial coefficients)는 다르지만 움직임 파라미터는 같은 영역들이 존재할 수 있다. 그러므로 부호화 관점에서 이러한 중복된 영역의 제거가 필요하다. 이것이 움직임 기준 영역 병합(motion-based region fusion)에 해당하며 영상 분할을 단순화하여 분할된 영역의 윤곽선 및 움직임 파라미터 정보를 크게 줄인다. 따라서 3.3절에서 분할된 영역들이, 움직임이 유사하고 서로 이웃해 있는 경우 이 영역들에 대해 병합을 행하게 되는데 이때 사용하는 움직임의 유사성은 affine 모형을 사용하여 테스트한다.

이웃한 두 영역의 병합 결정은 두 영역에 대해 최소 자승 오차 방법(least-squares technique)을 사용하여 구한 최적 affine 변환 계수와 그 오차 척도(error measure)를 이용하여 결정하게 된다. 최적 affine 변환 및 오차 척도를 구하는 방법은 다음과 같다.

• 최적 affine 변환의 계산:

영역 R에 대한 이동벡터(dense motion field)로부터 최소 자승 오차 방식을 이용하여 최적의 affine 변환계수를 계산한다. 이때 최소화해야 할 오차는 다음과 같다.

$$E(a_1, \dots, a_6) = \sum_R [(d_x(x, y) - a_1 - a_2x - a_3y)^2 + (d_y(x, y) - a_4 - a_5x - a_6y)^2] \quad (8)$$

여기서 $(d_x(x, y), d_y(x, y))$ 는 영역 R의 화소 (x, y) 에서 추정된 이동벡터이다. 따라서 이 오차 함수를 변환계수 $\{a_1, \dots, a_6\}$ 에 대하여 편미분한 뒤 그 식을 0으로 두면 6개의 선형 방정식이 얻어진다. 이

방정식이 상호독립이라면 그 해 $\{a_1^*, \dots, a_n^*\}$ 가 최적 affine 변환계수가 된다.

• 오차 척도(error measure):

후보영역 R_j 의 이동벡터와 계산된 최적 변환계수 $\{a_1^*, \dots, a_n^*\}$ 를 식 (8)에 대입하여 식 (9)와 같은 오차 척도 $\sigma(R_j)$ 를 계산한다.

$$\sigma(R_j) = \sqrt{\frac{E_j(a_1^*, \dots, a_n^*)}{\sum_{R_j}}} \quad (9)$$

$\sigma(R_j)$ 는 영역 $R_i \cup R_j$ 에 대한 최적 변환으로부터 후보영역 R_j 의 이동벡터 분포의 표준편차를 의미한다. 따라서 병합 결정은 $\sigma(R_j)$ 를 기준으로 하게 된다.

앞에서 언급한 최적 affine 변환 및 오차 함수를 이용하여 움직임이 유사한 이웃 영역을 병합하게 되는데 그림 3에 그 블록도를 나타내었다. 이 영역 병합법 [10]은 움직임 기준의 관점에서 균일한 작은 영역을 더 큰 영역으로 병합해 가는 영역 병합 과정으로 자세한 내용은 다음과 같다.

1. 병합을 위한 기준영역(seed region)을 찾는다. 아직 새로운 영역으로 할당되지 않은 영역들 중 가장 큰 영역을 기준영역으로 한다. 균일한 움직임을 갖는 영역 중 가장 큰 영역을 기준영역으로 선택하기 때문에 신뢰성이 큰 파라미터를 추정할 수 있다.
2. 기준영역 R_i 로 병합할 후보영역 R_j 를 찾는다. 여기서 후보영역(merging candidates)은 기준영역에 이웃한 영역 중 아직 새로운 영역으로 할당되지 않은 영역으로 한다.
3. 영역 $R_i \cup R_j$ 에 해당하는 최적 affine 변환 계수 $\{a_1^*, \dots, a_n^*\}$ 를 구한다.
4. 후보영역 R_j 의 병합결정을 위한 오차척도 $\sigma(R_j)$ 를 계산한다. 여기서 $\sigma(R_j)$ 는 영역 $R_i \cup R_j$ 에 대한 최적 변환으로부터의 R_j 에 있는 움직임 벡터의 표준편차를 의미한다.
5. 후보영역 R_j 을 기준영역 R_i 으로 병합할 것인지를 테스트한다. 후보영역 R_j 을 기준영역 R_i 로 병합할 것인지는 오차척도 $\sigma(R_j)$ 를 사용하여 결정한다. 만일 $\sigma(R_j)$ 가 주어진 임계값보다 적으면 후보 영역 R_j 는 기준 영역 R_i 로 병합된다. 기준영역은 후보영

역을 병합함에 따라 점차 영역이 확대된다. 따라서 추정된 변환계수는 점점 신뢰성이 높아진다.

6. 기준영역 주위에 더 이상 병합할 후보영역이 없을 때까지 단계 2-5을 반복적으로(recursively) 수행하게 된다. 기준영역 주위에 더 이상 병합할 후보가 없으면 새로운 기준영역을 찾아 병합을 수행하는 단계 1-5를 반복한다.

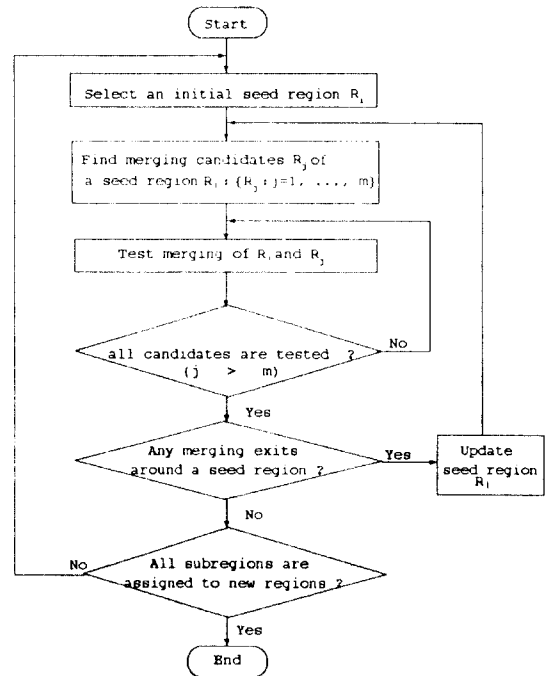


그림 3. 움직임 기준 영역병합 방법의 흐름도

IV. 모의실험 결과 및 고찰

본 절에서는 제안된 시공간 영상 분할 알고리즘을 이용하여 모의 실험한 결과를 제시하고 그 결과를 분석하였다. 모의실험에 사용된 테스트 시퀀스는 "Miss America" 영상(그림 4(a))과 "Carphone" 영상(그림 4(b))으로 QCIF(176×144)의 크기를 가지며 8 비트로 양자화되어 있다. 그리고 시공간 영상 분할을 위해 화소 단위의 이동 벡터장(dense motion field)을 이용하는데 참고문헌 [20]의 방법을 사용하여 추정하였다. 그림 4에 대한 이동 벡터장을 그림 5에 나타내었는데

이 이동 벡터장은 시간적으로 3 프레임 표본화된 프레임간에 추정된 것이다. 제안된 분할 알고리즘의 각 단계별 역할 및 중요성을 보이기 위해 아래에 각단계에 대한 결과를 나타내었다.

4.1 시공간 결합 marker 추출

시공간 결합 marker 추출을 위해 morphological open-close by reconstruction filter를 사용하여 먼저

영상을 단순화한다. 그림 5를 단순화한 영상을 그림 6에 나타내었다. 이때 사용된 structuring element 크기는 15×15으로 하였는데 그 이유는 머리-어깨(head and shoulder) 및 QCIF 해상도를 갖는 영상에 대해서는 그 크기가 경험적으로 적합했기 때문이다.

단순화된 영상으로 부터 intensity marker를 추출하는데 여기서 밝기값이 균일한 영역중 크기가 200 보다 큰 영역을 intensity marker로 추출하였다. 추출된



그림 4. 원영상: (a) "Miss America"(83th frame); (b) "Carphone"(106th) frame

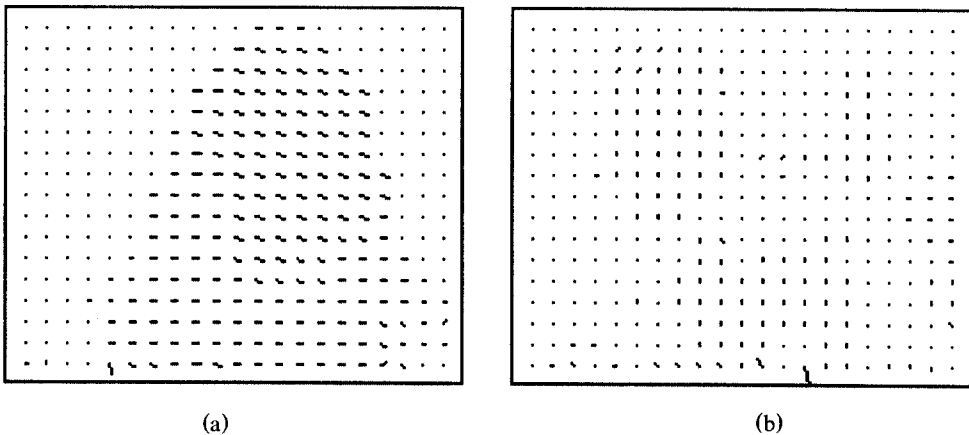


그림 5. 이동벡터장: (a) "Miss America"(between 80th and 83th frame); (b) "Carphone"(between 103th and 106th frame)

intensity marker를 그림 7에 나타내었다. 그림에서 각 marker의 라벨을 gray 값으로 표시하였고 joint marker가 아닌 영역, 즉 불확실한 영역(uncertain area)은 흰색으로 표시하였다. 그림에서도 알 수 있듯이 밝기만을 marker 추출에 이용할 경우 밝기가 비슷한 물체와 배경이 하나의 marker로 병합되어 나타나는 것을 알 수 있다. Intensity marker를 추출한 후, intensity

marker내의 균일한 움직임 영역을 joint marker로 추출하는데 그 결과를 그림 8에 표시하였다. 여기서 균일한 움직임은 affine 모형의 관점에서 추출하며 추출 방법은 그림 3에 나타낸 multipass 알고리즘을 사용하였다. 그림에서도 알 수 있듯이 시공간 정보를 이용한 결합 marker가 시각적으로 의미 있는 marker를 제공함을 알 수 있다.

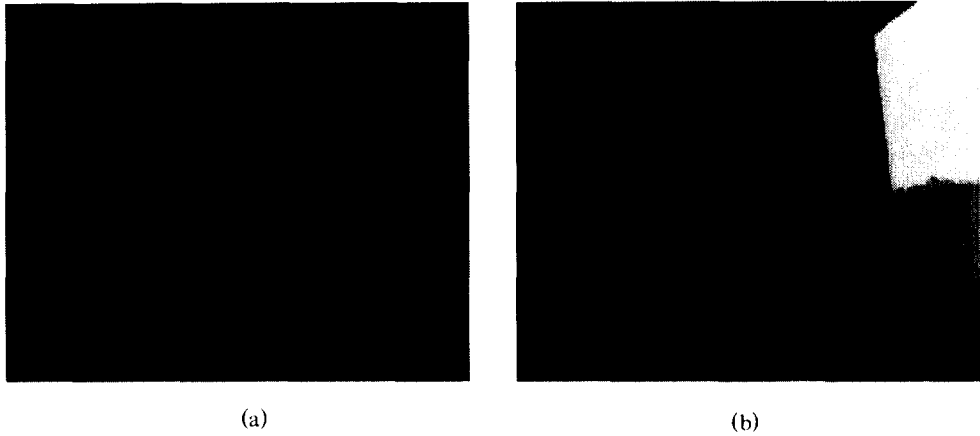


그림 6. 단순화된 영상: (a) "Miss America"(83th frame); (b) "Carphone"(106th frame)

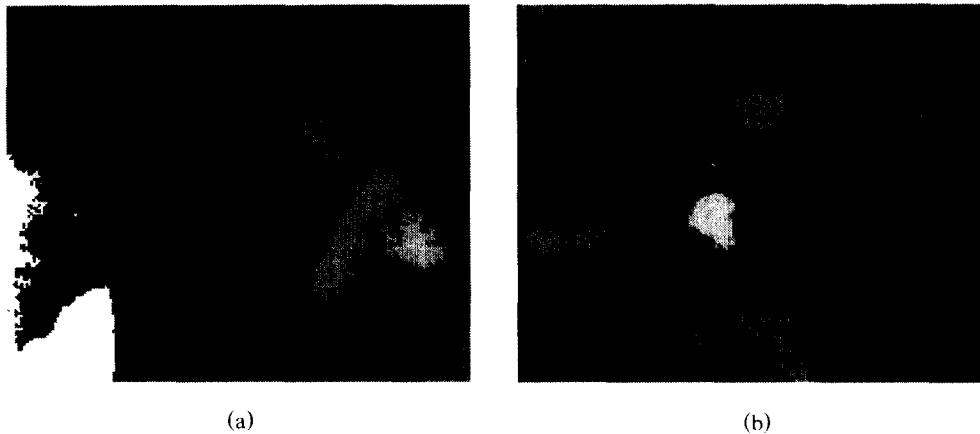


그림 7. Intensity markers: (a) "Miss America"; (b) "Carphone"



그림 8. 결합(joint) markers:(a) “Miss America”;(b) “Carphone”

4.2 영역 경계 결정

joint marker를 추출한 후 watershed 알고리즘에 의해 영역의 경계가 결정된다. 이때 식 (7)에 정의된 시공간 결합 척도(joint similarity measure)가 사용된다. 여기서 가중치 α 는 0.5로 하였다. 제안된 시공간 영상 분할의 결과를 morphology에 의한 공간 분할(morphological spatio segmentation) 결과와 비교하였는데 두 결과를 각각 그림 9와 그림 10에 나타내었다. 그림 9에 나타난 것처럼 공간 정보만에 의한 영상 분할은

밝기가 유사한 경우 배경(background) 및 물체(foreground objects)의 일부가 병합되어 분할됨을 알 수 있다. 그러나 제안된 시공간 영상 분할에서는 밝기가 유사한 물체의 경우에도 이들 사이에 정확한 영역 경계를 분할함을 알 수 있다. 이것은 결합 marker 추출과 시공간 결합 유사성 척도에 기인하는데 결합 marker가 시각적으로 의미있는 기준 영역을 제공하고 시공간 결합 척도가 움직임 물체에 대한 정확한 영역 경계를 제공하기 때문이다.

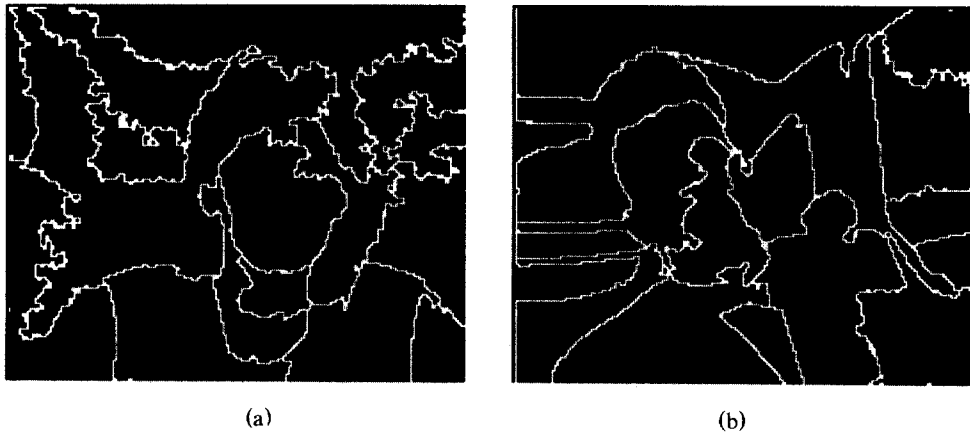


그림 9. Morphological 공간 분할:(a) “Miss America”;(b) “Carphone”

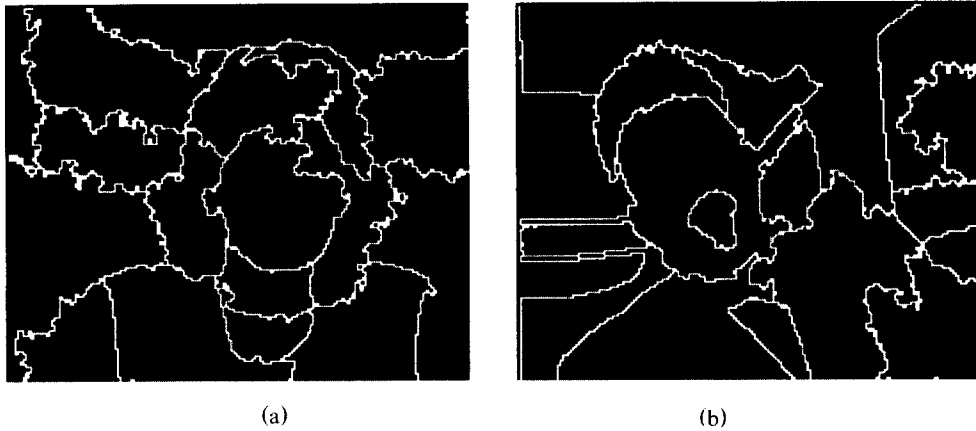


그림 10. 제안된 시공간 영상분할:(a) "Miss America";(b) "Carphone"

4.3 움직임 기준 영역 병합

그림 10에서 알 수 있듯이 앞절에서의 분할결과는 부호화 관점에서 중복성(redundancy)이 존재한다. 따라서 분할된 영역이 움직임이 유사한 경우 이를 병합하게 된다. 그 결과를 그림 11에 나타내었다. 그림 4(a)의 "Miss America"의 경우 배경과 머리-어깨 물체(foreground object)의 두 영역이 존재한다. 그림 11(a)에 주어진 분할 결과에서는 3 개의 영역으로 분할되었는데 배경 전체가 하나의 영역으로 분할되었고 머

리-어깨 물체가 얼굴과 나머지 물체의 두영역으로 분할되었다. 머리-어깨 물체의 경우 전체가 하나의 affine 변환으로 기술될 수 없기 때문에 두개의 영역으로 표현된 것이다. 한편 그림 4(b)에 주어진 "Carphone"의 경우 4 개의 물체로 구성되어 있는데 그것은 사람(human), 차의 내부 배경(car room) 그리고 2 개의 차밖 배경(outside background)이다. 그림 11(b)에서 알 수 있듯이 차의 내부 배경이 두개의 영역으로 분할되었는데 그 이유는 사람의 머리 부분이 움직이면서 그

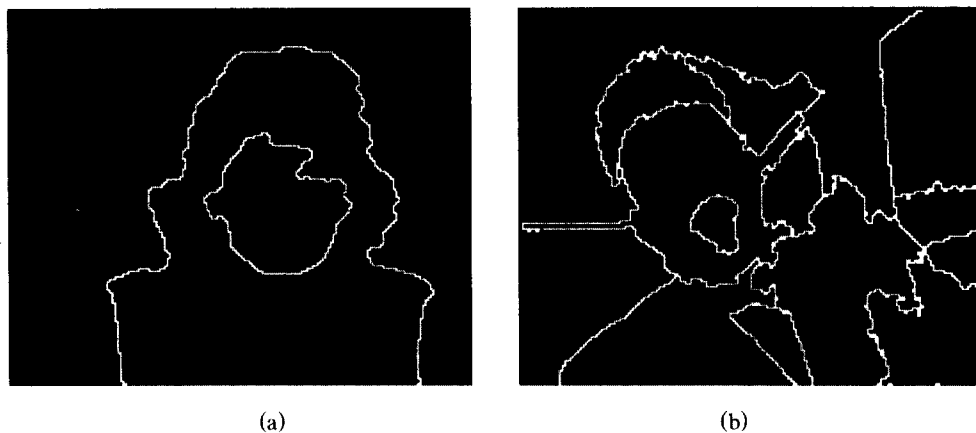
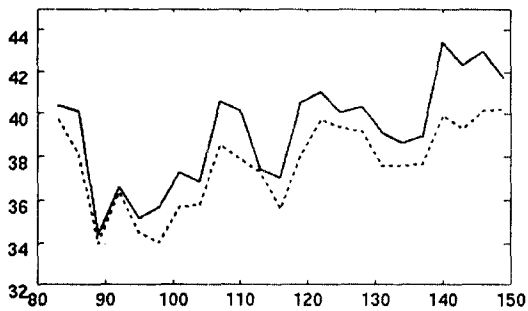


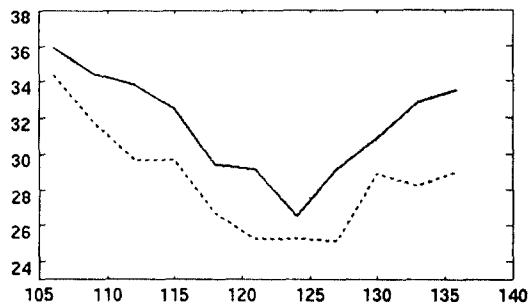
그림 11. 움직임 기준 영역 병합의 결과:(a) "Miss America";(b) "Carphone"

림자로 인해 차의 내부 배경의 밝기가 변했기 때문이다. 그리고 사람 및 차밖 배경이 각각 7 개 및 4 개의 영역으로 분할되었다. 실험 결과에서 알 수 있듯이 제안된 시공간 영상 분할이 시각적으로 의미있는 결과를 제공함을 알 수 있다.

제안된 영상 분할 방법을 영상 부호화에 적용할 경우 부호화 효율을 보이기 위해 제안된 시공간 영상 분할 방법을 이용하여 분할된 결과에 대해 영역별로 움직임 추정 및 보상을 행한 결과를 그림 12에 나타내었다. 움직임 보상 결과의 성능을 비교하기 위하여, 참고문헌 [12]의 시공간 영상분할 방법에 의한 움직임 보상 결과도 함께 도시하였다. 움직임 보상 결과의 성능 평가는 PSNR로 하였다. 그림 12(a)에 "Miss America" 영상 시퀀스에 대한 PSNR을 도시하였고 그림 12(b)에 "Carphone" 영상 시퀀스에 대한 PSNR을 도시하였다.



(a)



(b)

그림 12. 영역별로 움직임을 보성한 결과(PSNR):(a) "Miss America"; (b) "Carphone"

그림에서도 알 수 있듯이 제안된 분할 방법에 의한 움직임이 보상 결과가 전반적으로 우수한 PSNR을 나타냄을 알 수 있다. 이것은 제안된 분할 방법이 식 (7)의 결합 유사성 척도를 사용하여 움직임 보상 오차가 최소가 되도록 영역을 분할하기 때문이다. 따라서 제안된 분할 방법이 물체별(object-based) 또는 영역기반(region-based) 부호화에 적용될 경우 높은 부호화 성능을 나타낼 수 있을 것이다.

V. 결 론

본 논문에서는 효율적인 시공간 영상 분할 알고리즘을 제안하였다. 제안 방법은 밝기 정보와 움직임 정보를 동시에 고려하고 morphological filter와 watershed 알고리즘을 기본 도구(tool)로서 사용한다. 전체적인 분할 과정은 3 단계로 구성되는데 그것은 joint marker 추출, 영역 경계 결정, 움직임 기준 영역 병합이다. 시각적으로 의미있는 marker 추출을 위해 간단하고 효과적인 joint marker 추출 알고리즘을 제안하였다. 그리고 밝기값과 움직임을 동시에 고려하는 영역 분할을 위해 새로운 결합 척도(joint similarity measure)를 제안하였다. 또 움직임 기준 영역 병합을 위해 효과적인 영역 병합법을 사용하였다. 실험 결과에서 주어진 것처럼 제안된 시공간 영상 분할 알고리즘이 시각적으로 의미있는 분할 결과를 제공함을 알 수 있고 정확한 영역 경계를 결정하기 때문에 영상 부호화 적용시 영역 경계에서 특히 우수한 재생(reconstruction) 화질을 나타낼 것이다. 또한 제안된 방식은 사전 지식을 필요로 하지 않는다는 점에서 자동적인 영상 분할(automatic segmentation)을 가능케 한다.

앞으로의 연구는 두가지 방향으로 이루어 질 것이다. 첫째, 연속된 프레임의 분할 결과 사이에 유사성(segmentation similarity)이 유지되도록 하기 위한 time-recursive 분할에 대한 연구가 진행될 것이다. 이러한 연속된 프레임 분할 결과 사이의 유사성은 물체별 또는 영역기반 부호화에서 윤곽선 정보(contour parameters)를 줄이는데 매우 중요하기 때문이다. 둘째, 제안된 시공간 분할 방법을 영상 부호화에 응용하는 것이 연구될 것이다.

참 고 문 헌

1. T. Ebrahimi, E. Reusens and W. Li, "New trends in very low bitrate video coding," *Proc. IEEE*, vol. 83, no. 6, pp. 877-891, June 1995.
2. M. Kunt, A. Ikonomopoulos, and M. Kocher, "Recent results in high compression image coding," *IEEE Trans. CAS*, vol. 34, pp. 1306-1336, Nov. 1987.
3. H. G. Musmann, M. Hotter and J. Osterman, "Object oriented analysis-synthesis coding of moving images," *Signal Processing: Image Communication*, vol. 1, no. 2, pp. 117-138, Oct. 1989.
4. P. Salembier, L. Torres, F. Meyer and C. Gu, "Region-based video coding using mathematical morphology," *Proc. IEEE*, vol. 83, no. 6, June 1995, pp. 843-857, June 1995.
5. D. Cortez, P. Nunes, M. Sequeira, and F. Pereira, "Image segmentation towards new image representation methods," *Signal Processing: Image Communication*, vol. 6, no. 6, pp. 485-498, Feb. 1995.
6. P. Salembier, "Morphological multiscale segmentation for image coding," *Signal Processing*, vol. 38, pp. 359-386, 1994.
7. P. Salembier and M. Pardas, "Hierarchical morphological segmentation for image sequence coding," *IEEE Trans. on Image Processing*, vol. 3, no. 5, pp. 639-651, Sep. 1994.
8. F. Dufaux and F. Moscheni, "Motion estimation techniques for digital TV: a review and a new contribution," *Proc. IEEE*, vol. 83, no. 6, pp. 858-876, June 1995.
9. M. Hotter and R. Thoma, "Image segmentation based on object oriented mapping parameter estimation," *Signal Process.*, vol. 15, no. 3, pp. 315-334, Oct. 1988.
10. J. G. Choi, S. W. Lee and S. D. Kim, "Segmentation and motion estimation of moving objects for object-oriented coding," *Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing*, Detroit, MI, May 9-12, 1995, pp. 2431-2434.
11. G. Adiv, "Determining three-dimensional motion and structure from optical flow generated by several moving objects," *IEEE Trans. on PAMI*, vol. 7, no. 4, pp. 384-401, July 1985.
12. J. Konrad, A. Mansouri, E. Dubois and V. Dang, "On motion modeling and estimation for very low bit rate video coding," *Proc. Visual Communications and Image Processing '95*, vol. 2501, pp. 262-273, 1995.
13. L. Wu, J. Benois-Pineau, D. Barba, "Region-based time-varying image coding at low bit-rate with a high visual quality," *Proc. Visual Communications and Image Processing '94*, vol. 2308, pp. 1074-1085, 1994.
14. N. T. Watsuji, H. Katata and T. Aono, "Morphological segmentation with motion based feature extraction," *International workshop on coding techniques for very low bit-rate video*, Tokyo, Nov. 8-10, 1995.
15. C. Gu, T. Ebrahimi and M. Kunt, "Morphological spatio-temporal segmentation for content-based video coding," *International workshop on coding techniques for very low bit-rate video*, Tokyo, Nov. 8-10, 1995.
16. W. H. Hong, N. C. Kim, and S. M. Lee, "Video segmentation using spatial proximity, color, and motion information for region-based coding," *Proc. Visual Communications and Image Processing '94*, vol. 2308, 1994, pp. 1627-1633.
17. D. H. Ballard and C. M. Brown, *Computer vision*, Prentice-Hall, Englewood Cliffs, New Jersey, 1982, Ch. 5, pp. 151-152.
18. 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.
19. B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, pp. 185-203, 1981.
20. J. H. Lee and S. D. Kim, "Velocity field estimation using a weighted local optimization,"

IEICE Trans. Fundamentals, vol. E76-A, no. 4, pp. 661-663, April 1993.

21. H. H. Nagel, "Displacement vectors derived from second-order intensity variation in image sequences," *Comput. Vision Graphic Image Processing*, vol. 21, pp. 85-117, 1993.
22. J. H. Lee and S. D. Kim, "An Error analysis of gradient-based methods," *Signal Processing*, vol. 35, no. 2, pp. 157-162, 1994.
23. P. Mulroy, "Spatial and temporal image segmentation by feature clustering," *International workshop on coding techniques for very low bit-rate video*, Tokyo, Nov. 8-10, 1995.



崔在覺(Jae Gark Choi) 정회원
 1961년 5월 9일생
 1980년 3월~1984년 2월: 경북대학교 전자공학과 졸업(공학사)
 1985년 3월~1987년 2월: 한국과학기술원 전기 및 전자공학과 졸업(공학석사)

석사)

1992년 3월~현재: 한국과학기술원 재학중

※주관심분야: 영상처리, 영상통신, VLSI 구현 등



李時雄(Si-Woong Lee) 정회원
 1969년 1월 10일생
 1987년 3월~1991년 2월: 경북대학교 전자공학과 졸업(공학사)
 1991년 3월~1993년 2월: 한국과학기술원 전기 및 전자공학과 졸업(공학석사)

석사)

1993년 3월~현재: 한국과학기술원 재학중

※주관심분야: 영상처리, 영상통신, VLSI 구현 등



趙順濟(Soon-Jae Cho) 정회원
 1971년 2월 15일생
 1988년 3월~1992년 2월: 경북대학교 전자공학과 졸업(공학사)
 1992년 3월~1994년 2월: 한국과학기술원 전기 및 전자공학과 졸업(공학석사)

석사)

1994년 3월~현재: 한국과학기술원 재학중

※주관심분야: 영상처리, 컴퓨터 비전, VLSI 구현 등



金聖大(Seong-Dae Kim) 정회원
 1953년 12월 26일생
 1977년: 서울대학교 전자공학과 졸업(공학사)
 1979년: 한국과학기술원 전기 및 전자공학과 졸업(공학석사)
 1983년: 프랑스 INPT ENSEEIHT

졸업(공학박사)

1984년~현재: 한국과학기술원 전기 및 전자공학과 교수

※주관심분야: 영상처리, 영상통신, 컴퓨터 비전, VLSI 구현 등