

이진분열 알고리즘에 기반한 계층적 구조의 시공간 영상 분할

박 영 식, 송 근 원, 정 의 윤, 한 규 필, 하 영 호
 경북대학교 전자전기공학부

E-mail : yspark@paigong.kyungpook.ac.kr Fax: 053-957-1194

Spatio-Temporal Image Segmentation Using Hierarchical Structure Based on Binary Split Algorithm

Young Sik Park, Kun Woen Song, Eui Yoon Chung, Kyu Phil Han and Yeong Ho Ha
 School of Electronic & Electrical Eng. Kyungpook Nat'l Univ.

Abstract

In this paper, a hierarchical spatio-temporal image segmentation method based on binary split algorithm is proposed. Intensity and displacement vector at each pixel are used for image segmentation. The displacement vectors between two image frames which skip over one or several frames can be approximated by accumulating of the velocity vectors calculated from optical flow between two successive frames when the time interval between the two image frames is short enough or the motion is slow.

The pixels whose displacement vector and intensity are ambiguous are precisely decided by the modified watershed algorithm using the proposed priority measure. In the experiment, the region of moving object is precisely segmented.

분할기반 부호화 기법에서는 영상을 움직임(motion), 밝기값(intensity), 그리고 색차신호 등을 기준으로 연결성과 동질성을 만족하는 영역들로 분할한 뒤 각 영역의 경계를 표현하는 윤곽선(contour)과 영역 내부의 질감(texture)을 분리하여 부호화 한다. 일반적으로 윤곽선 정보는 인간시각 인식에 있어 질감정보 보다 중요할 뿐만 아니라 부호화에 소요되는 비트 량도 질감정보 보다 많은 비중을 차지한다. 따라서 낮은 비트율 환경에서는 최소의 윤곽선 정보를 이용하여 영상을 효율적으로 표현하는 분할기법이 필요하다.

집합이론에 기반한 비선형적 특징을 나타내는 수리 형태학(mathematical morphology) 필터는 영상에서 물체에 대한 형태, 크기, 대조 등의 다양한 특징을 추출할 수 있고 동질성과 연결성을 만족하는 부분에 대한 특징 추출이 효과적이기 때문에 영상분할에 많이 이용되고 있다[1-3].

1. 서 론

멀티미디어와 통신망 기술의 개발로 PSTN(public switched telephone network), LAN(local area network), ATM(asynchronous transfer mode), 그리고 이동통신망을 통한 낮은 비트율 환경에서의 영상정보의 서비스에 대한 요구가 증가되고 있다. H.261과 MPEG-1과 같은 블록단위의 움직임 보상과 DCT를 이용하는 부호화 기법을 낮은 비트율 환경에 그대로 적용하면 블록화(blocking) 현상과 모스키토(mosquito) 현상에 의한 화질의 열화가 두드러진다. 또한 멀티미디어 응용에서 요구되는 객체단위의 다양한 처리가 요구되어진다. 이러한 문제점을 해결하기 위한 차세대 부호화 기법의 표준화가 MPEG-4와 ITU-T에서 진행중이며 여러 부호화 기법 중에서 객체기반 부호화(object-based coding) 또는 분할기반 부호화(segmentation-based coding) 기법이 활발히 연구되고 있다.

II. 이진분열에 의한 계층적인 시공간 영상분할

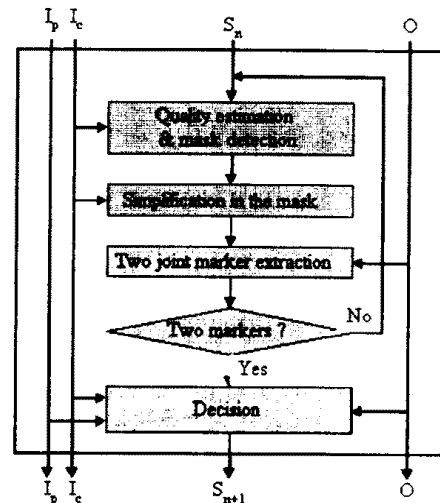


그림 1. 이진분열에 기반한 시공간 영상 분할을 위한 제안된 계층적 구조

그림 1의 제안된 이진 분할에 의한 계층적 시공간 영상 분할 기법은 광류(optical flow)를 이용하여 추출한 변위벡터와 밝기값을 영역을 분할하는 척도로 이용하여 움직임 영역을 정확히 분할한다. 여기서 O 는 광류에 의하여 추출된 변위벡터이고 I_c 와 I_p 는 각각 현재 영상과 이전 영상을 의미하고 S_n 은 n 개의 영역으로 분할된 영상이다. 첫 번째 계층에서 입력으로 사용되는 분할영상은 영역의 개수가 하나인 원영상의 평균값으로 사상된 영상을 이용한다.

1. 광류(optical flow)를 이용한 변위벡터 추정

광류는 밝기값 패턴의 움직임에 대한 속도의 분포를 나타내며 추정하는 영상간의 시간 간격이 짧고 움직임이 적은 경우에 속도(velocity) 벡터는 변위(displacement) 벡터로 근사화된다. 따라서 그림 2에서와 같이 현재 영상 I_p 와 이전 영상 I_c 사이의 변위벡터는 인접한 두영상에서 구한 각각의 속도벡터의 합으로 근사화되어질 수 있다.

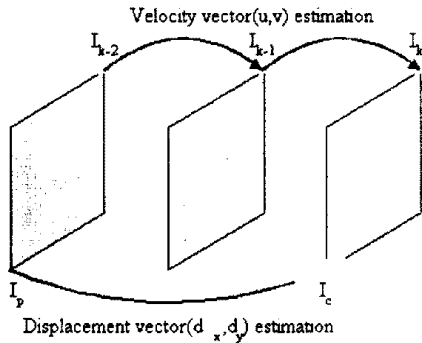


그림 2. 광류를 이용한 변위벡터의 추정

이때 속도벡터 (u, v) 를 계산하기 위해서 (x, y, t) 에서의 밝기값과 $(x + \delta x, y + \delta y, t + \delta t)$ 에서의 밝기값이 같다는 가정하에서

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t) \quad (1)$$

와 같으며 우측항은 Taylor 급수에 의하여 전개되면

$$I(x, y, t) = I(x, y, t) + \delta x \frac{\partial I}{\partial x} + \delta y \frac{\partial I}{\partial y} + \delta t \frac{\partial I}{\partial t} + \varepsilon \quad (2)$$

와 같이 표현되어진다. 여기서 고차항은 작은 값이므로 무시하고 δt 에 대하여 양변을 나누고 이를 정리하면

$$\frac{\delta x}{\delta t} \frac{\partial I}{\partial x} + \frac{\delta y}{\delta t} \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} = 0 \quad (3)$$

와 같다. 식 (3)은 속도벡터와 밝기값의 기울기 성분의 관계를 나타내는 것으로

$$G_x u + G_y v + G_t = 0 \quad (4)$$

와 같이 다시 표현되어질 수 있다. 그러나 식 (4)에서는 유일한 해인 속도벡터 (u, v) 를 찾지 못하기 때문에 인접한 속도벡터는 연속성(smoothness)을 가진다는 조건을 도입하여 유일한 해를 결정한다. 여기서 밝기값 변화율에 대한 오차는

$$E_i^2 = (G_x u + G_y v + G_t)^2 \quad (5)$$

와 같이 표현되고 연속성 조건에 의한 오차는

$$E_c^2 = \left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 + \left(\frac{\partial v}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial y} \right)^2 \quad (6)$$

와 같이 표현 되어진다. 이를 이용하여 전체 오차는

$$E^2 = \iint (E_i^2 + \alpha^2 E_c^2) dx dy \quad (7)$$

와 같이 표현되어지며 이 오차를 최소화하는 속도벡터 (u, v) 를 찾으면 된다. 반복기법에 의한 해는

$$\begin{aligned} u^{n+1} &= \bar{u}^n - G_x \{ G_x \bar{u}^n + G_y \bar{v}^n + G_t \} / \{ \alpha^2 + G_x^2 + G_y^2 \} \\ v^{n+1} &= \bar{v}^n - G_y \{ G_x \bar{u}^n + G_y \bar{v}^n + G_t \} / \{ \alpha^2 + G_x^2 + G_y^2 \} \end{aligned} \quad (8)$$

에 의하여 구할 수 있다. 여기서

$$\begin{aligned} \bar{u}_{i,j,k} &= \frac{1}{6} \{ u_{i-1,j,k} + u_{i+1,j,k} + u_{i,j-1,k} + u_{i,j+1,k} \} \\ &\quad + \frac{1}{12} \{ u_{i-1,j-1,k} + u_{i-1,j+1,k} + u_{i+1,j-1,k} + u_{i+1,j+1,k} \} \\ \bar{v}_{i,j,k} &= \frac{1}{6} \{ v_{i-1,j,k} + v_{i+1,j,k} + v_{i,j-1,k} + v_{i,j+1,k} \} \\ &\quad + \frac{1}{12} \{ v_{i-1,j-1,k} + v_{i-1,j+1,k} + v_{i+1,j-1,k} + v_{i+1,j+1,k} \} \end{aligned} \quad (9)$$

와 같으며

$$\begin{aligned} G_x &\approx \frac{1}{4} \{ I_{i,j-1,k} - I_{i,j,k} + I_{i+1,j+1,k} - I_{i+1,j,k} \\ &\quad + I_{i,j+1,k+1} - I_{i,j,k+1} + I_{i+1,j+1,k+1} - I_{i+1,j,k+1} \} \\ G_y &\approx \frac{1}{4} \{ I_{i+1,j,k} - I_{i,j,k} + I_{i+1,j+1,k} - I_{i,j+1,k} \\ &\quad + I_{i+1,j,k+1} - I_{i,j,k+1} + I_{i+1,j+1,k+1} - I_{i,j+1,k+1} \} \end{aligned}$$

$$G_t \approx \frac{1}{4} \{ I_{i,j,k+1} - I_{i,j,k} + I_{i+1,j,k+1} - I_{i+1,j,k} + I_{i,j+1,k+1} - I_{i,j+1,k} + I_{i+1,j+1,k+1} - I_{i+1,j+1,k} \} \quad (10)$$

와 같이 구한다.

2. 화질추정 및 마스크 검출

각각의 영역에 평균값으로 사상된 분할 영상에서 화질이 저하된 영역을 검출하여 미세한 영역으로 분할함으로써 화질의 개선을 가져올 수 있다. 분할된 영역들 중에서 다음 계층에서 두 개의 영역으로 분할되는 마스크를 검출하기 위한 척도로는 원 영상과 분할영상의 평균 절대차(mean absolute difference)를 화질을 평가하기 위한 척도로 이용한다.

3. 마스크내에서의 단순화

모폴로지컬 필터를 이용한 단순화 필터는 재구성에 의한 열림과 닫힘 필터(open close by reconstruction)가 주로 사용되어진다[1,2]. 본 논문에서는 연산량을 줄이기 위하여 작은 크기의 형태소 하나만을 이용하여 마스크내에서 단순화를 한다. 이는

$$\begin{aligned} \delta_n^M(f)(x) &= \max \{ f(x-y), y \in S_n \text{ and } y \in M \}, \\ \epsilon_n^M(f)(x) &= \min \{ f(x+y), y \in S_n \text{ and } y \in M \} \end{aligned} \quad (11)$$

와 같이 변형된 불림(dilation)과 부식(erosion)을 사용하며 여기서 M 은 마스크를 의미한다. 그리고 이를 이용한 재구성에 의한 열림과 닫힘 필터는

$$\varphi_M^{(rec)}(\delta_n^M(\gamma_M^{(rec)}(\epsilon_n^M(f), f)), \gamma_M^{(rec)}(\epsilon_n^M(f), f)) \quad (12)$$

와 같이 표현된다. 이는 마스크내에서 단순화를 수행하게 된다.

4. 두 개의 마커추출

기존의 분할방법에서 각 계층별로 추출되는 마커의 개수에 의하여 영역의 개수가 결정되기 때문에 발생하는 정보량을 조절하기가 매우 어렵다. 따라서 주어진 비트율에서 최대의 화질을 얻기 위해서는 영역의 개수를 조절하면서 분할하는 방법이 필수적이다.

제안된 이진 분열에 의한 분할은 각 계층에서 MAD가 가장 낮은 하나의 영역을 두 개의 영역으로 분할하기 때문에 영역의 개수를 조절할 수 있다. 이는 단순화된 영상에서 단순화 필터의 형태소 면적보다 큰 동일한 영역은 마커로 추출한다. 여기서 동일한 영역이란 평균밝기 값과 변위벡터가 동일한 영역을 의미한다. 추출되는 여러 개의 마커중에서 두 개의 마커를 선택하는 것이 분할된 영상의 화질을 좌우하는데 제안한 알고리즘에서는 밝기 값

이 가장 큰 것과 가장 작은 두 개의 마커를 선택한다. 이는 하나의 영역을 대조가 가장 큰 두 개의 영역으로 분리하여 하나의 영역 증가로 많은 주관적인 화질의 개선을 가져오기 위한 것이다. 또한 단순화 필터의 형태소 크기에 분할의 결과가 좌우되지 않기 때문에 작은 크기의 형태소를 이용하여 단순화함으로써 연산량을 줄일 수 있고 하드웨어 복잡도가 줄어든다. 만약 두 개 이상의 마커가 추출되지 않으면 그 영역에서는 분할을 중단하고 남은 영역 중에서 MAD가 가장 낮은 영역에 대하여 위의 과정을 반복한다. 분할을 중단하는 것은 영역의 면적이 작아서 분할하더라도 더 이상의 화질개선이 없다고 판단되는 영역이다.

5. 변형된 분수계(watershed) 알고리즘에 의한 영역결정

마스크 내에서 두 개의 마커를 중심으로 변형된 계층적인 분수계 알고리즘을 이용하여 불확실한 영역을 결정함으로써 두 개의 영역으로 분할한다. 이는 동일한 영역의 초기 라벨링 과정과 불확실한 화소들의 범람 과정으로 이루어진다.

1) 동일한 영역의 초기 라벨링

추출된 두 개의 마커는 분할의 시발점으로 마스크 내에서 동일한 영역, 즉 밝기값과 변위벡터가 동일한 영역으로 분할된 영역의 내부에 포함되는 영역이다. 따라서 마킹된 영역은 어떤 영역에 속할지 확실한 영역이다. 이것의 라벨링 과정은 영상을 스캔(scan)하면서 가장 먼저 발견되는 라벨링 되지않은 마커의 한 화소의 좌표값을 큐에 삽입한다. 그리고 큐에서 한 화소를 추출하여 '1'을 라벨링을 한다. 그리고 이웃의 4화소 중에서 화소의 밝기값이 같고 라벨링이 되어있지 않으면 그 화소의 좌표값을 큐에 삽입한다. 이러한 과정을 큐가 빌 때까지 반복하면 하나의 마커에 대하여 라벨링이 끝나게 된다. 그리고 라벨값을 1만큼 증가시키고 위의 과정을 반복하면 모든 마커된 영역에 대하여 라벨링이 끝나게 된다.

2) 불확실한 화소들의 범람

동일한 영역의 초기 라벨링 단계에서 라벨링된 영역에 인접한 모든 불확실한 화소들을 그림 3에서와 같이 우선순위에 따라 큐에 삽입한다. 이때 우선순위를 결정하는 거리는 큐에 삽입되는 화소의 밝기 값과 인접영역의 평균 밝기값 그리고 인접영역의 평균 변위벡터에 의한 이전영상 I_p 에서 추정된 밝기값과의 각각의 차에 가중치를 곱한 합으로

$$distance = \alpha |I_c(R) - I_c(x, y)| + (1 - \alpha) |I_p(x - d_{R_x}, y - d_{R_y}) - I_c(x, y)| \quad (13)$$

와 같이 표현된다. 여기서 큐에 삽입되는 화소의 변위벡터가 아니라 인접영역의 평균 변위벡터를 이용하여 이전영상 I_p 에서 추정된

밝기값과 큐에 삽입되는 화소의 밝기 값의 차를 이용하여 거리를 구한 이유는 경계에서 정확한 영역을 분할하기 위한 것이다. 실험에서 가중치 α 는 0.25를 사용하였다. 그리고 우선순위는

$$priority = \frac{1}{distance} \quad (14)$$

와 같이 거리에 반비례한다.

가장 높은 우선순위의 큐에서 한 화소를 추출하고 이 화소가 아직 라벨링이 되지 않은 화소이면 이 화소와 거리가 가장 작은 영역의 라벨을 추출한 화소의 위치에 부여한다. 그리고 인접한 4 화소 중에서 라벨링이 되지 않은 화소는 우선순위를 계산하여 큐에 삽입한다. 이때 인접화소들이 이미 큐에 삽입되어 있는 경우에는 우선 순위가 높은 경우에만 다시 큐에 삽입한다. 추출된 화소가 이미 라벨링된 화소라면 인접화소들 중에서 아직 라벨링이 되지 않은 화소들을 우선순위에 의해 큐에 삽입한다.

추출된 화소에 대하여 라벨링이 되면 그 라벨에 대한 영역의 평균 밝기값과 변위벡터는 갱신된다. 이와같은 과정을 큐가 빌 때까지 되풀이 하게되면 모든 화소가 어떤 영역에 속하게 되는지 결정되며 분할을 마치게 된다.

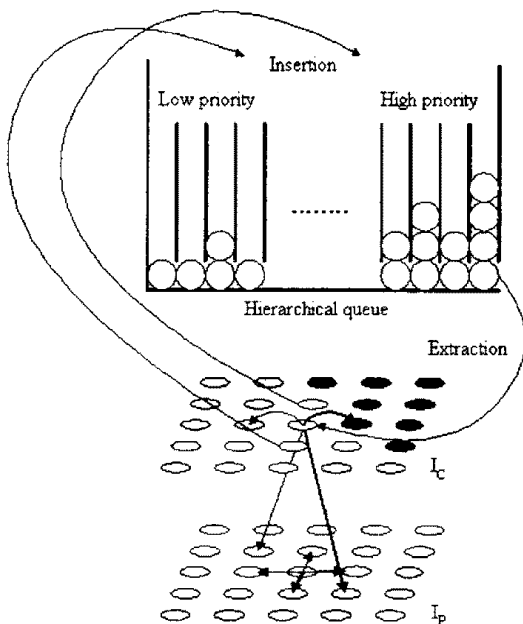


그림 3. 변형된 분수계 알고리즘에서 시공간 거리를 이용한 결정

III. 실험결과 및 고찰

실험은 176×144 QCIF 영상에 대하여 모의실험을 하였다. 그림 4에서는 원영상을 그림 5에서는 제안된 이진분열에 의한 계층적인 시공간 영상분할 결과와 경계선 보여주며 이때의 영역의 개수는 150개이다. 그림 6에서는 변위벡터에 의한 움직임 영역을 자동분할한 결과를 보여준다. (a)에서는 경계선을 단순화하기 전이고 (b)에서는 경계선 단순화 기법[5]을 이용하여 한 화소 두개의 잡음영역을 제거한 결과이다. 그림 7의 (a)에서는 광류에 의하여 추정된 변위벡터를 나타낸다. 이때 영역의 경계선이 정확하지 않으나 제안된 시공간 거리에 의한 우선순위에 의한 분할후의 영역별 변위벡터(b)에서는 움직임 영역의 경계를 정확히 분할하는 것을 볼 수 있다.

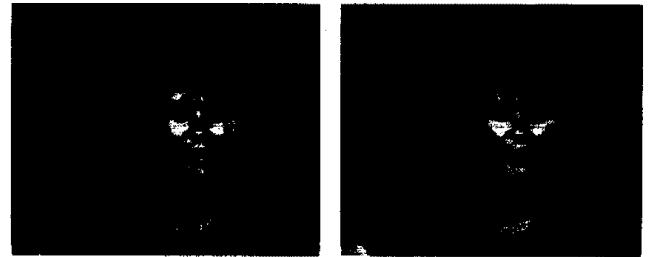


그림 4. 실험에 사용된 79번째와 81번째 Miss America 영상

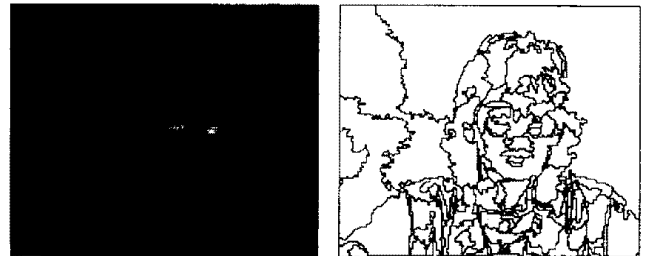


그림 5. 이진분열에 의한 계층적인 시공간 분할영상과 경계선

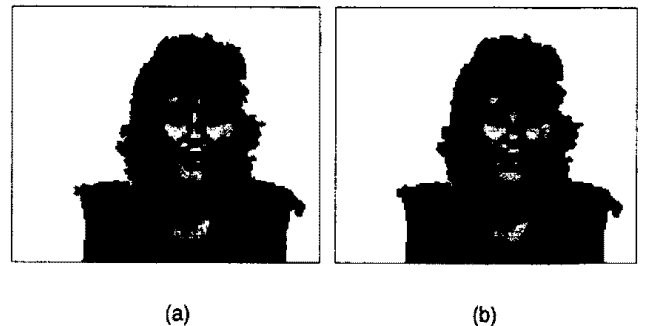
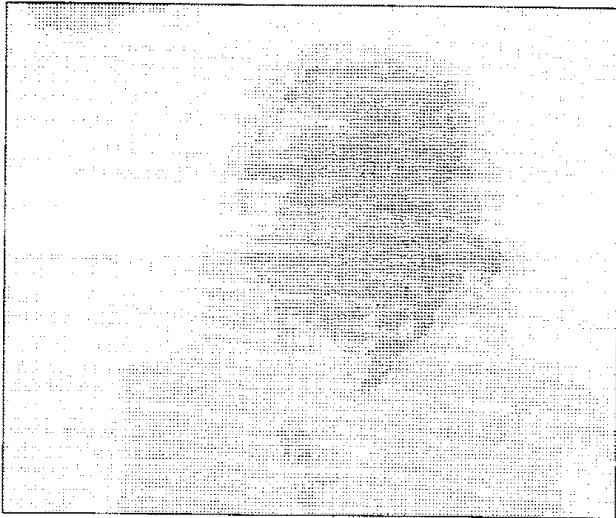
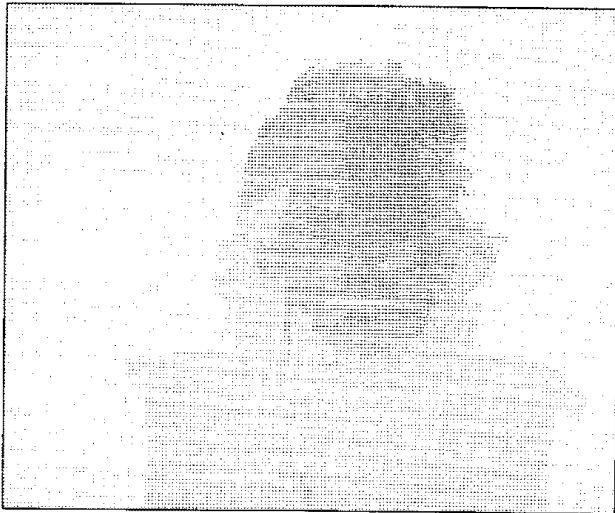


그림 6. 움직임 영역의 자동분할 결과 (a) 경계선 단순화 전 (b) 경계선 단순화 후



(a)



(b)

그림 7. 변위벡터 (a) 광류에 의한 변위벡터 (b) 분할후의 각 영역에 대한 변위벡터

IV. 결 론

본 논문에서 제안된 이진 분열을 이용한 계층적 시공간 영상 분할 알고리즘은 화질이 가장 낮은 하나의 영역에서 변위벡터가 동일하고 대조가 가장 큰 두 개의 마커를 추출하여 두 개의 영역으로 분리한다. 이 과정에서 작은 크기의 하나의 형태소에 의한 형태학 필터를 이용하여 단순화하기 때문에 기존의 계층적인 분할 알고리즘에 비하여 연산량을 크게 줄일 수 있고 구조가 간단하다. 대조가 가장 큰 두 개의 영역으로 분열하기 때문에 발생하는 영역의 개수를 조절할 수 있어 적은 개수의 영역으로 효율적으로 영상을 표현할 수 있어 낮은 비트율 환경에서 사용하기에 적합하다. 그리고 광류를 이용하여 변위 벡터를 추정하고 이를 밝기 값과 함께 영역을 분할하는 기준으로 이용함으로써 움직임 영

역을 정확하게 분할할 수 있다.

참 고 문 헌

- [1] P. Salembier, "Morphological multiscale segmentation for image coding," *Signal Processing*, vol. 38, pp. 359-386, 1994.
- [2] 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.
- [3] J. G. Choi, S. W. Lee, and S. D. 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.
- [4] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, pp. 185-203, 1981.
- [5] 박영식, 김기석, 송근원, 정의운, 하영호, "디지털 이동통신망에서의 초저속 영상부호화를 위한 영역단위의 계층적 분할과 경계선 단순화 기법," *한국통신학회논문지*, 제22권, 3호, pp. 432-443, 1997년 3월.