

시공간 엔트로피 임계법을 이용한 형태학적 이동 객체 분할

백경환*, 신민수**, 곽노윤*

*천안대학교 정보통신학부

**한국전자통신연구원, 위성멀티미디어연구팀

Moving Object Segmentation Using Spatio-temporal Entropic Thresholding

Kyoung-Hwan Baek*, Min-Su Shin**, No-Yoon Kwak*

*Division of Information and Communication Engineering, Cheonan University

**Satellite Multimedia Research Team, ETRI

요 약

본 논문은 비디오 시퀀스에 카메라 패닝 보상과 2차원 시공간 엔트로피 임계법을 적용하여 추출한 객체포함영역을 대상으로 영상 분할을 수행하는 이동 객체 분할 기법에 관한 것이다. 우선, 웨이블릿 변환에 의해 구성된 피라미드 계층 구조상에서 카메라 패닝 벡터를 추정하여 전역 움직임 보상을 한다. 이후, 전역 움직임이 보상된 기준영상을 대상으로 각 프레임간에서 2차원 시공간 엔트로피 임계법을 적용하여 이동 객체가 포함될 가능성이 있는 영역을 블록 단위로 추출한다. 다음으로, 2차원 시공간 엔트로피 임계법에 의해 분류된 영역을 토대로 각 블록을 움직임블록, 준 움직임 블록, 비 움직임 블록 중 어느 하나로 분류한 검색 테이블을 작성한다. 이어, 검색 테이블을 참조하여 초기 탐색 계층 및 탐색 영역을 적응적으로 선정함으로써 피라미드 계층 구조상에서 효율적인 고속 움직임 추정을 수행하여 이동 객체에 해당하는 객체포함영역을 추출한다. 최종적으로, 이렇게 추출된 객체포함영역에서 임계 기술기 영상을 정의한 후, 이를 기준 삼아 객체포함영역에 화소 단위의 형태학 기반 영상 분할 알고리즘을 적용함으로써 비디오 시퀀스에 포함된 이동 객체를 분할한다. 컴퓨터 시뮬레이션 결과를 통해 고찰할 때, 제안된 방법은 이동 객체에 대한 상대적으로 우수한 분할 특성을 제공할 수 있고, 특히 저대조 경계면의 분할 특성을 제고시키고 있음을 확인할 수 있다.

1. 서 론

최근 들어 MPEG-4로 대표되는 내용 기반 영상 코덱의 적용 범위가 날로 확장되고 있고 이와 연계된 국제 표준들이 점차적으로 완성·실용 단계에 접어들면서 영상 분할 기술은 이 분야의 실용화를 위한 핵심 요소 기술로 크게 부각되고 있다. 특히, 내용 기반의 기능성을 제공하는 다양한 응용 분야의 출현이 강하게 요청되면서 이를 제공할 시에 가장 중요한 기초가 되는 비디오 객체(video object)를 추출하는 알고리즘이 다각도로 연구되고 있다.

기존의 컬러나 밝기값과 같은 영상 신호가 가지는 공간 정보(spatial information)의 통계적 또는 형태학적인 특성만을 이용하는 방법[1][2]의 경우, 분할 영역을 의미론적으로 해석하고자 할 시에 많은 제한과 한계를 드러내고 있는 것이 사실이다. 이에 따라 객체는 배경 영역과 다른 움직임을 갖는다는 가정 하에 움직임을 이용한 영상 분할 기법이 제안되어 있다. 하지만 움직임 정보는 잡음에 민감하여 불확실한 결과를 제공할 수 있고 움직임 객체의 경계선에서 배경이 가려지거나 또는 새롭게 드러나기 때문에 비디오 객체 추출을 위한 정확한 윤곽선 검출이 어려운 난제가 있다.

이러한 문제를 극복하기 위해 컬러나 밝기값과 같은 정보와 함께 움직임 정보를 같이 사용하여 비디오 객체를 추출하는 시공간 영상 분할 기법이 다수 제안되고 있다[3][4]. 이러한 시공간 영상 분할 기법은 밝기값과 움직임 정보를 함께 이용한다는 관점에서 움직임이나 밝기값 하나만을 이용하는 영상 분할 기법보다 우수한 결과를 제공하지만, 밝기값 또는 움직임을 이용하여 각각 독립적으로 영상을 분할한 후, 이 결과를 조합하는 방식으로 영상 분할이 이루어지기 때문에 밝기값이 비슷한 물체가 하나의 물체로 병합되는 경우 이를 수정하기가 어려운 단점이 있다.

제안된 방법은 이러한 단점을 보완하기 위해 초기에 전역 움직임 추정을 통해 카메라의 움직임에 의한 영향을 제거한 상태에서 2차원 시공간 엔트로피 임계법(2D spatio-temporal entropic thresholding)을 적용하여 이동 객체가 포함될 가능성이 있는 후보 영역을 블록 단위로 추출한다. 이후, 이렇게 추출된 후보 영역을 대상으로 HBMA(Hierarchical Block Matching Algorithm)를 적용하여 이동 객체에 해당하는 영역만을 추출한 다음에 임계 기술기 영상을 기준 삼아 형태학 기반 영상 분할 기법을 적용함으로써 이동 객체

에 대한 상대적으로 우수한 분할 특성을 제공하고자 한다.

2. 기존의 2차원 엔트로피 임계법

T. Pun에 의해 처음 소개된 1차원 엔트로피 임계법(1D entropic thresholding)은 배경과 객체 영역의 정보량을 최대화함으로써 영상의 모든 화소들을 객체 클래스와 배경 클래스로 분할하는 최적의 임계치를 결정한다[5]. 이에 기초하여 Fan et al.은 영상 프레임의 시간적 공간적 상관 관계를 동시에 고려할 수 있는 2차원 임계값 벡터로서, 그림 1과 같이 프레임차대비(FCON; Frame Difference Contrast)와 국부분산대비(LCON; Local Variance Contrast)로 구성된 분할 평면에서 최적의 임계 벡터를 구하고, 이를 이용하여 영상 시퀀스에서 이동 객체를 시공간적으로 분할하는 방법을 제안하고 있다[6]. 이는 프레임간 움직임 보상성(compensability)을 향상시키는 물론 영상 내에 포함되어 있는 더 많은 정보를 사용하기 때문에 결국에는 좋은 분할 결과를 제공한다.

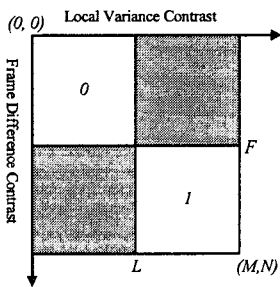


그림 1. 기존의 2차원 엔트로피 임계법의 분할 평면

그림 2는 기존의 2차원 엔트로피 임계법을 적용한 분류 영상을 예시한 것으로, FCON과 LCON을 사용해 구한 최적의 임계 벡터 (\bar{F} , \bar{L})를 기준 삼아 변화 영역, 변화가 없는 배경과 객체 및 불확실 영역으로 분류한 중간 결과 영상이다.



그림 2. Bike#40에 2차원 엔트로피 임계법을 적용한 분류 영상

이 알고리즘에서는 영상의 시간적인 정보를 사용하기 위해서 프레임차만을 이용하였기 때문에 배경이 복잡한 영상에서 전역 움직임이 존재하는 경우에 변화가 없는 영역이 움직임이 있는 객체나 배경으로 분할되는 경우가 발생하게 된다. 초기에 잘못 분할되는 이러한 블록들은 다시 복원되기가 어렵게 되어 결국 영상 내의 객체를 원형 그대로 분할하는데 실패할 확

률이 높다는 단점이 있다.

3. 제안된 이동 객체 분할

제안된 방법은 이러한 단점을 보완하기 위해 제안된 것으로, 크게 네 부분으로 구성되어 있다. 첫 번째는 전역 움직임 추정을 통해 카메라의 움직임에 의한 영향을 제거하는 과정이고, 두 번째는 전역 움직임이 제거된 상태에서 2차원 시공간 임계법을 적용하여 이동 객체가 포함될 가능성이 있는 후보 영역을 추출하는 과정이고, 세 번째는 이렇게 추출된 영역을 대상으로 HBMA를 적용하여 이동 객체에 해당하는 영역만을 추출하는 과정이며, 네 번째는 이동 객체 영역을 대상으로 임계 기술기 영상을 기준 삼아 형태학 기반 영상 분할 기법을 적용함으로써 이동 객체의 형상 정보를 추출하는 영상 분할 과정이다.

3.1 계층적 전역 움직임 보상

제안된 전역 움직임 보상은 웨이블릿 변환을 수행함에 따라 발생하는 각 계층의 저주파 부영상만을 이용하여 3 단계의 계층으로 구성된 피라미드 계층 공간상에서 이루어진다. 피라미드 계층 공간의 최상위 계층에서 최하위 계층 방향으로 진행하면서 점진적으로 카메라 패닝 벡터의 추정 정밀도를 재조정하여 최종적으로 카메라 패닝 벡터를 추정한 후, 이를 이용해 전역 움직임을 보상함으로써 카메라의 움직임에 의한 영향을 제거하는 과정으로 이루어진다.

연속된 실제 영상에서 카메라 패닝은 시험 영상에 따라 다르겠지만, 확률적으로 그 발생 빈도수가 낮기 때문에 카메라 패닝을 항상 고려하여 보상하는 것은 과도한 연산량을 수반하게 됨에 따라 전체적으로 움직임 추정 시간을 지연시키는 비효율적 요인이 된다. 따라서 적은 연산량을 필요로 하면서도 카메라 패닝의 발생 여부를 효과적으로 감지할 수 있는 판정 기준이 필요하다. 이를 위해, 제안된 전역 움직임 보상 방식에서는 최하위 계층 영상에서 차영상 윤곽정보를 이용하여 정의한 평균 블록 활동도가 실험적으로 정한 임계값 보다 클 경우에 한해서 카메라 패닝이 발생한 것으로 판단하고 전역 움직임 보상을 위한 연산을 수행한다.

카메라 패닝이 발생한 것으로 판단될 경우, 우선, 탐색 공간에 대한 여유를 확보하기 위해 카메라 패닝 추정 블록을 최하위 계층 영상을 기준으로 할 때, 좌우 상하 ± 16 pixels를 들여서 정한다. 이후, 각 계층의 해상도에 맞게 조정된 이 블록을 가지고 피라미드 계층 공간의 최상위 계층 영상에서 ± 1 pixel, 중간 계층 영상에서도 ± 1 pixel, 그리고 마지막으로, 최하위 계층 영상에서도 ± 1 pixel의 탐색 영역을 고려하여 카메라 패닝 벡터를 계층적으로 추정한다. 이에 따르면, 원 영상을 기본 탐색 공간으로 가정할 때, 카메라 패닝 벡터의 탐색 영역을 ± 7 pixels로 설정한 것과 동일한 탐색 범위를 갖는데, 계층적인 추정 방식임에 따라 추정 속도를 고속화할 수 있다. 여기서, 탐색 영역을 ± 7 pixels로 고려한 것은 통상의 카메라 패닝은 한 프레임 사이에서 7개 화소 이상 이동하는 경우의 발생 빈도수가 매우 낮음을 실험적으로 확인한 바에 따른 것이며, 설사, ± 7 pixels 이상의 카메라 패닝이 발생한다고 하더라도 이 경우에는 화면에 극도의 뭉뚱화

(blurring) 현상이 발생됨에 따라 카메라 패닝을 보상하는 행위 자체가 무의미하기 때문이다.

3.2 2차원 시공간 엔트로피 임계법

제안된 엔트로피 임계법은 전역 움직임이 보상된 비디오 시퀀스에서 전역움직임보상 프레임차대비(GCON; Globally Motion-Compensated Frame Difference Contrast)와 국부분산대비(LCON; Local Variance Contrast)로 구성된 2차원 분할 평면에서 최적의 임계 벡터를 구하고, 이를 이용하여 영상 시퀀스에서 이동 객체가 포함된 후보 영역을 추출한다.

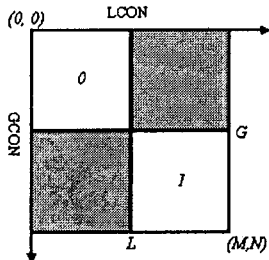


그림 3. 제안된 2차원 엔트로피 임계법의 분할평면

2차원 엔트로피 평면을 구성하는 GCON과 LCON은 각각 식 (2)와 식(3)과 같이 계산된다.

$$GCON(m) = \sum_{x=0}^{L-1} \sum_{y=0}^{G-1} |I(x, y, t_n) - I(x-d_x, y-d_y, t_p)|^2 \quad (2)$$

여기서, $I(x, y, t_n)$ 은 현 프레임의 m 번째 블록 내의 (x, y) 위치의 화소값이고, $I(x-d_x, y-d_y, t_p)$ 는 전역 움직임이 보상된 참조 영상 내의 화소값이며, d_x 와 d_y 은 카메라 패닝 벡터이다.

$$LCON(m) = \sum_{x=0}^{L-1} \sum_{y=0}^{G-1} |I(x, y, t_n) - \bar{I}(m)|^2 \quad (3)$$

여기서, $\bar{I} = (1/255) \sum_{x=0}^{L-1} \sum_{y=0}^{G-1} I(x, y, t_n)$ 는 현 블록의 평균 화소값이다.

한편, 배경 및 객체 가능 영역에 해당하는 두 클래스의 엔트로피는 각각 $H_B(G, L)$ 과 $H_O(G, L)$ 로 나타내며 다음과 같이 구할 수 있다.

$$H_B(G, L) = - \sum_{i=0}^{G-1} \sum_{j=0}^{L-1} P_{B_{ij}}(G, L) \log P_{B_{ij}}(G, L) \quad (4)$$

$$H_O(G, L) = - \sum_{i=G+1}^{M-1} \sum_{j=L+1}^{N-1} P_{O_{ij}}(G, L) \log P_{O_{ij}}(G, L) \quad (5)$$

영상 분할 수행을 위한 최적의 전역 임계치 벡터 (\bar{G}, \bar{L}) 는 다음과 같은 기준 함수를 만족해야 한다.

$$H(\bar{G}, \bar{L}) = \max \left(\min_{\substack{G=0..M-1 \\ L=0..N-1}} \{H_B(G, L), H_O(G, L)\} \right) \quad (6)$$

이와 같이 구해진 전역 임계치 벡터 (\bar{G}, \bar{L}) 를 기준으로 조건식 (7)을 통해 움직이는 객체, 배경 및 불확

실 영역으로 분류한다.

$$\begin{aligned} GCON(m) > \bar{F}, LCON(m) < \bar{L} & \text{ changed region} \\ GCON(m) < \bar{F}, LCON(m) < \bar{L} & \text{ stationary background} \\ GCON(m) < \bar{F}, LCON(m) > \bar{L} & \text{ stationary object} \\ GCON(m) > \bar{F}, LCON(m) > \bar{L} & \text{ uncertain region} \end{aligned} \quad (7)$$

식 (7)에 의해 객체, 배경, 불확실 블록으로 분류된 결과를 토대로 객체에 해당하는 블록을 움직임 블록, 불확실 영역에 해당하는 블록을 준 움직임 블록, 배경 영역에 해당하는 블록을 비 움직임 블록으로 분류한 검색 테이블(LUT; Look Up Table)을 작성한다.

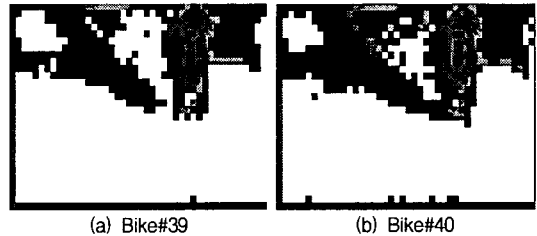


그림 4. Bike#39와 Bike#40에 대해 제안된 2차원 시공간 엔트로피 임계법을 적용한 블록 분류 영상

3.3 HBMA을 이용한 블록 분류

제안된 HBMA는 전역 움직임 보상 과정에서 사용한 바 있는 3 단계의 피라미드 계층 공간을 이용한다. 움직임 추정을 위한 기본단위는 최하위 계층 영상에서의 추정 블록의 크기는 16pixels×16pixels 매크로 블록으로 정하였으며, 이에 따라 중간 계층 영상에서의 추정 블록의 크기는 8pixels×8pixels, 최상위 계층 영상에서는 4pixels×4pixels로, 계층이 증가함에 따라 추정 블록의 크기는 옥타브(octave)적으로 감소한다.

우선, 2차원 엔트로피 임계법에 의해 분류된 블록들을 토대로 작성된 LUT을 이용하여 움직임 추정을 수행한다. LUT상에서 비 움직임 블록으로 분류된 블록은 움직임 벡터가 0일 것으로 가정하고, 움직임 탐색을 수행하지 않는 상태에서 영벡터를 할당함에 따라 불필요한 움직임 추정 과정을 생략함으로써 움직임 추정 속도의 고속화를 도모한다. 한편, 준 움직임 블록으로 분류된 블록은 초기 탐색 계층을 최상위 계층으로 선택하고, 초기 탐색 계층으로 선택한 최상위 계층에서 ±3pixels, 중간 계층에서 ±2pixels, 최하위 계층에서 ±1pixel의 순으로 탐색 영역을 설정하여 움직임 벡터를 검출한다. 마지막으로, LUT상에서 움직임 블록으로 분류된 블록은 국부 최소에 빠질 확률을 줄이기 위해 초기 탐색 계층을 중간 계층 영상으로 선택하고, 이 계층에서 ±7pixels, 다음으로 최하위 계층에서 ±4pixels의 탐색 영역을 설정하여 움직임 벡터를 검출한다.

이렇게 추정된 움직임 벡터의 동질성을 이용하여 그림 5와 같은 이동 객체 블록들을 추출한다. 움직임 벡터의 동질성에 기반하여 블록 분류를 수행하면 객체 이외의 배경 부분이 대부분 제거되지만, 그림 5에서 볼 수 있듯이, 배경과 객체 영역에 고립된 작은 블록들이 잔존하게 된다.

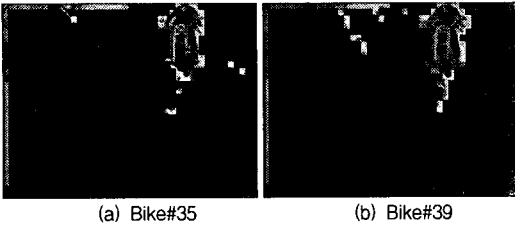


그림 5. 움직임 벡터를 이용한 블록 분류 기법을 적용한 영상

이러한 고립 블록을 제외시키기 위해 연결 성분 분석(connected-component analysis) 기법을 적용하여 가장 큰 성분 요소를 객체로 간주하고 나머지 성분들은 모두 제거함으로써 그림 6과 같은 객체포함영역을 추출한다.

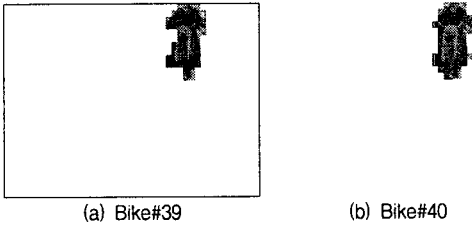


그림 6. 제안된 방법에 의해 추출된 객체포함영역

3.4 형태학 기반 영상 분할

이상과 같이 블록 단위로 추출된 객체포함영역에 대해 형태학적 영상 분할 기법을 적용하여 동질성과 연속성을 만족하는 화소 단위의 영역들로 분할한다. 본고에서는 저대조 평탄 영역에서도 객체의 윤곽선을 보존하기 위해 임계 기울기 영상(thresholded gradient image)를 이용하여 평탄 영역과 비평탄 영역으로 분할 후 각 영역별로 마커 추출 영역을 제한하여 영상을 분할한다. 임계 기울기 영상을 기준 삼아 객체포함영역을 평탄 영역과 비평탄 영역으로 분리함으로써 저대조 경계면의 분할특성을 제고시킬 수 있다.

그림 7은 임계 기울기 영상의 추출 과정을 나타낸 것이다. 영상의 기울기를 구하기 위해 식 (8)과 같이 정의되는 다중척도 형태학적 기울기 연산자(multiscale gradient operator)[7]를 사용한다.

$$MG(f_s) = \frac{1}{n} \sum_{i=1}^n [((f_s \oplus B_i) - (f_s \ominus B_i)) \ominus B_{i-1}], \quad (8)$$

for $1 \leq i \leq n$

여기서, B_i 는 $(2i+1) \times (2i+1)$ 척도의 정사각형 형태소를 나타낸 것이다. 이후, 이 기울기 영상은 기 설정된 임계치에 따라 평탄 영역과 비평탄 영역으로 구분하는 기준이 되고 이를 참조하여 각 영역은 독립적으로 마커를 추출한다.

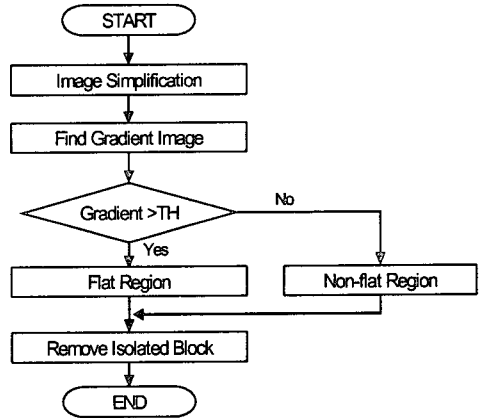


그림 7. 임계 기울기 영상의 추출 과정

이후의 마커 추출 단계는 다음과 같이 구성된다. 우선, 임계 기울기 영상에 의해 제한된 영역 내에서 상호 인접하는 동일 화소값들을 갖는 그룹들에 같은 레이블을 붙이고 서로 인접하지 않은 다른 값으로 레이블링한다. 다음으로, 각각의 레이블링된 이웃하는 화소 그룹들의 크기가 임계치보다 크면 마커로서 지정이 되고 그렇지 않은 경우 불확실한 미결정 화소로서 남게 된다.

마커 추출 단계를 수행한 후에는 아직 어느 영역에도 지정되지 않은 미결정 화소들이 남아 있다. 이 화소들을 어느 한 영역에 지정하기 위해 그림 8에 나타난 변형된 워터셰드 알고리즘을 이용함으로써 그림 9에 나타낸 바와 같은 이동 객체를 분할할 수 있다.

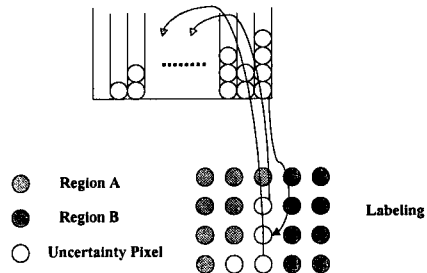


그림 8. 변형된 워터셰드 알고리즘

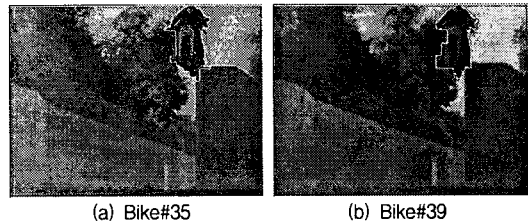


그림 9. 제안된 방법을 통해 추출한 이동 객체 영상

4. 결론

제안된 방법은 비디오 시퀀스에 카메라 패닝 보상과 2차원 시공간 엔트로피 임계법을 적용하여 추출한 객체포함영역을 대상으로 임계 기울기 영상을 이용한 형태학 기반 영상 분할을 수행함으로써 이동 객체에 대한 분할 특성의 향상을 기할 수 있었다.

컴퓨터 시뮬레이션 결과를 통해 고찰할 때, 제안한 방법은 카메라 패닝 보상을 통해 전역 움직임을 보상함으로써 전역 움직임에 의해 배경 영역이 객체포함 영역으로 오인될 가능성을 효과적으로 경감시키고 있음을 알 수 있다. 다시 말해서, 제안된 방법이 채택한 GCON은 FCON의 경우와 비교해 볼 때 배경 영역에 해당하는 부분을 더욱 효과적으로 제거한 것을 알 수 있다. FCON을 사용할 경우, 전역 움직임을 가진 배경 블록들이 움직임이 있는 것으로 나타나 객체 블록이나 불확정 블록의 수가 많았으나, GCON을 사용하여 전역 움직임을 보상한 경우에는 이러한 블록들이 모두 배경 블록으로 분류되어 결과적으로 객체 블록과 불확정 블록의 수가 경감되고 배경 블록의 수가 늘어난 것을 알 수 있다.

또한, 2차원 시공간 엔트로피 임계법에 의해 배경 영역과 객체포함영역이 적절하게 분리되고 있으며, 형태학 기반 영상 분할 알고리즘을 적용할 시에 임계 기울기 영상을 참조함으로써 저대조 경계면에서 분할 특성을 제고시킬 수 있음을 확인할 수 있었다.

하지만, 제안된 방법은 영상 내에 존재하는 객체나 객체의 일부가 매우 작은 움직임만을 가지는 경우나 두 개 이상의 움직이는 객체가 떨어져서 존재하는 경우 등에 대해서는 적절히 고려되지 않았으므로 이 부분에 대한 추가적인 연구가 진행되어야 할 것으로 판단된다. 또한, 다양한 분포 특성을 갖는 영상에 대해 전역적인 파라미터를 적용할 수 있도록 그 성능을 개선함으로써 제안된 방법의 타당성과 보편성을 좀 더 폭넓게 확보할 필요가 있다.

[참고문헌]

- [1] 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.
- [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] M. 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, pp. 8-10, Nov. 1995.
- [4] Jae Gark Choi, Si Woong Lee and Seong Dae Kim, "Video segmentation based on spatial and temporal information," *Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing*, Munich, Germany, 20-24, Apr. 1997.
- [5] T. Pun, "Entropy thresholding : a new method,"

Comput. Vis. Graph. Image Process. 47, pp. 22-32, 1981.

- [6] J. Fan, L. Zhang, F. Gan, "Spatiotemporal segmentation based on two-dimensional spatiotemporal entropic thresholding," *Society of Photo-Optical Instrumentation Engineers*, pp. 2845-2851, Oct. 1997.
- [7] D. Wang, "A Multiscale Gradient Algorithm for Image Segmentation Using Watersheds", *Pattern Recognition*, vol. 30, no. 12, pp. 2043-2052, 1997.