계층적 모션 추정을 통한 장면 분할 기법

김 모 곤, 우 종 선, *정 순 기
선린대학 인터넷테크노&멀티미디어 계열, *경북대학교 컴퓨터공학과
전화 : 054-260-2209 / 핸드폰 : 011-9562-2489

Scene Segmentation using a Hierarchical Motion Estimation Technique

Mo Gon Kim, Jong Sun Woo, *Soon Ki Jung
Sunlin College, *Kyungpook National University
E-mail : mgkim@sunlin.ac.kr

Abstract

We propose the new algorithm for scene segmentation. The proposed system consists motion estimation module and motion segmentation module. The former estimates 2D-motion value for each pixel position from two images transformed by wavelet. The latter determine scene segments well fitting on dominant affine motion models. What distinguishes proposed algorithm from other methods is that it needs not other post-processing for scene segmentation. We can manipulate both multimedia data and objects in virtual environment using proposed algorithm.

Ⅰ. 서 론

현재 컴퓨터 그래픽스 분야에서 연구가 되어오고 있는 영상기반 모델링 및 렌더링(Image-Based Modeling and Rendering) 기법은 멀티미디어 데이터의 사실성을 높이는데 한 몫을 하고 있다.

본 논문에서는 동영상으로부터 추출된 영상 프레임들을 대상으로 모션 정보를 얻은 후, 유사한 모션 모델로 판단되는 물체들을 그룹화시켜 전체 영상 프레임을 분할하기 위한 영상 기반 모델링 기법을 소개한다. 이를 위하여 새로운 모션 추정 알고리즘과 모션 정보를 이용한 영상 분할 알고리즘을 제시한다.

특히, 모션 분할(motion segmentation) 기법은 컴퓨터비전 분야에서 다루고 있는 고전적인 문제라 할 수 있다. 일반적으로 모션 정보에 의한 영상 분할 기법의 결과는 추정된 모션의 정확도와 그 알고리즘에 많은 영향을 받는다. 따라서 본 논문에서는 다단계 웨이블릿 변환 영상으로부터 비교적 신뢰할 만한 모션 정보를 추출한 후, 모션 모델에 의한 영상 분할 작업 알고리즘을 기술한다.

이를 통해 멀티미디어 데이터의 조작 및 편집 등에 이용할 수 있는 가능성을 제시한다.

논문의 구성은 다음과 같다. 2장에서는 다단계 웨이블릿 변환 영상으로부터 모션을 추정하는 방법에 대해 기술하고, 3장에서는 모션 정보를 이용하여 모션 모델링을 통한 모션 분할 알고리즘을 자세하게 설명한다. 마지막으로 4장에서 결론 및 향후과제를 기술한다.

Ⅱ. 2D 모션 추정 알고리즘

2D 모션이라 함은 3차원 물체의 움직임을 영상 평면에 투영시킨 광 흐름(optical flow)을 일컫는다. 이러한 2D 모션의 추정은 컴퓨터 비전분야에서 오랫동안 연구되어 온 문제들 중 하나이지만, 컴퓨터비전의 특성상 제안된 방법들이 모든 응용분야에 적용되지는 않는다.

일반적으로 2D 모션은 시간의 변화에 따른 영상 내의 명암(intensity) 패턴으로부터 인식된다. 즉, 2D 모션은 영상 프레임에서의 밝기 값이 같은 대응 픽셀(correspondence)을 찾음으로써 구해지기 때문에, 영상 프레임 내에 조명의 변화 또는 잡음이 포함되어 있을 경우에 2D 모션 추정이 어려운 단점이 있다. 실제로 2D 모션 모델링에 있어서 모델의 파라미터 값은 2D 모션 추정 값의 정확도에 많은 영향을 받는다.

본 논문에서는 영상 프레임에서의 급격한 밝기 변화에 대한 근사를 잘 할 수 있는 웨이블릿[2-3]을 이용하여 영상을 변환시킨 뒤 저주파 영상만으로 2D 모션을 추정한다. 일반적으로 웨이블릿 변환 기법은 공간 영역과 주파수 영역에서의 지역적인 특성을 잘 나타내고 서로 다른 해상도로 영상을 분할하여 개별적으로 부분 영상을 분석할 수 있는 특성 때문에 다해상도 분석(multi-resolution analysis)에 널리 이용된다.

2D 모션을 추정하는 경우에도 영상 프레임 내에 포함되어 있는 잡음이나 물체의 경계선 부분에 대한 2D 모션을 비교적 정확하게 추정할 수 있게 된다. 그러나 웨이블릿 변환 영상에서 저주파 영상만을 이용하여 2D 모션을 추정하게 되면 물체의 경계선 부분에 대한 모션을 평활화(smoothing)시키는 결과를 가져올 수도 있기 때문에, 본 논문에서는 다단계 웨이블릿 변환으로 얻은 영상에 대해 모션을 단계적으로 추정하는 방법을 이용한다. 이렇게 2D 모션 추정 방법을 달리한 이유는 모션을 모델링하는 과정에서 잘못 추정된 모션 값으로 인해 모션 모델의 파라미터 값에 미칠 수 있는 영향을 최소화하기 위해서이다.

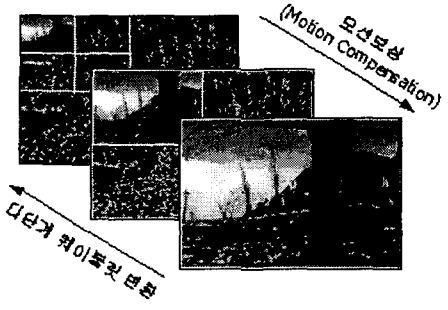
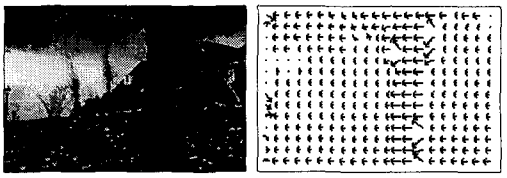


그림 1. 다단계 웨이블릿 변환

본 논문에서는 비교적 간단한 Haar 웨이블릿을 이용하여 변환된 저주파 영상을 대상으로 블록매칭을 수행하고, 그림 1과 같이 단계적으로 모션 값을 보상하면서 2D 모션을 추정(Coarse-to-Fine)한다. 그림 2는 추정된 2D 모션을 시각적으로 잘 보여주고 있다.



(가) 첫 번째 영상프레임 (나) 추정된 2D 모션

그림 2. 2D 모션을 추정한 결과

Ⅲ. 모션 분할(Motion Segmentation)

본 장에서는 2D 모션의 해석을 위해 어파인(affine) 모션 모델을 도입하고, 이를 통해 장면을 분할하기 위한 알고리즘을 기술한다.

일반적으로 영상 분할로 생성된 각각의 영역들은 어떤 동질성을 가지고 있다. 실제로 이러한 속성들이 영상 분할에서 아주 중요한 요소가 되고, 이러한 동질성을 .판단하기 위한 기준으로는 텍스쳐(texture), 흑백정도, 혹은 의미 있는 정보를 가진 모델 등이 될 수 있다. 모션 분할은 영상에서 같은 모션 값을 가지는 영상 내의 물체나 영역을 나타내는 픽셀들에 대해 레이블링을 수행하는 것이라 정의할 수 있다.

그러나 2D 모션 정보만으로는 모션 해석 방법에 한계가 있기 때문에, 본 논문에서는 어파인 모션 모델을 도입하여 2D 모션을 모델링한 후에, 장면(scene)을 분할하고자 한다.

모션을 모델링한다는 의미는 카메라의 움직임을 모델링하는 것과 일치하므로, 어파인 모션 모델은 카메라의 평행 이동시에 발생하는 시점 운동(motion parallax)을 비교적 잘 표현할 수 있다.

어파인 모델에 기반한 모션 분할의 대표적인 연구로는 적응적 어파인 모델 클러스터링(adaptive affine clustering) 기법[1]을 예로 들 수 있다. 그러나 이 방법은 다음과 같은 단점을 발견할 수 있다.

첫째, 모델이 가지는 어파인 모델 파라미터의 유사도를 $D_m(a_1, a_2) = [(a_1 - a_2)^T M (a_1 - a_2)]^{1/2}$ 으로 평가하여 결정하는데, 이것은 물리적인 근거가 없다. 즉, 가중치 행렬 M을 어떻게 설정하느냐에 따라 클러스터링 결과가 민감하다. 둘째, 어파인 모델 파라미터를 조금만 잘못 추정해도 장면 분할 결과에 영향을 미친다. 마지막으로, 파라미터 클러스터링과 레이블링 과정이 서로 분리되어 있기 때문에 원하는 장면 분할 결과를 위하여 후처리 과정이 필요하다.

이러한 이유로 Altunbasak[5]은 컬러정보를 이용하여 영역 기반 모션 세그멘테이션 기법을 제안하였고,

Borshukov[6] 등은 다단계 어파인 모델 분류 기법을 제안한 바 있다.

본 논문에서 제안한 모션 분할 시스템의 블록도를 그림 3에 제시하고, 시스템을 크게 초기 모션 모델 생성 단계와 모션 모델 클러스터링 단계로 나누어 기술한다.

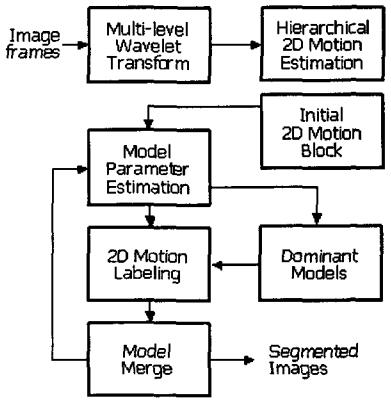


그림 3. 모션 분할 시스템 블록도

3.1 초기 모션 모델 결정

우선 입력 영상을 16×12의 블록 B_i로 나누어 각 블록을 하나의 모션 모델로 간주하고, 식(1)을 이용하여 어파인 모션 모델 파라미터를 추정한다.

$$u_a(x,y) = (a_{i1}-1)x + a_{i2}y + a_{i3}$$
$$v_a(x,y) = a_{i4}x + (a_{i5}-1)y + a_{i6} \quad \text{식(1)}$$

$u_a(x,y)$와 $v_a(x,y)$는 블록 B_i내에 있는 모든 픽셀 (x,y)에 대한 어파인 모션 값을 의미하며, 파라미터 $a_{i1}, a_{i2}, \cdots, a_{i6}$는 회귀분석법을 이용하여 추정할 수 있다. 회귀분석법은 독립변수와 종속변수 사이의 함수 관계를 데이터에 의해 규명하고자 할 때 사용되는 통계적 방법 중의 하나이다.

본 논문에서는 회귀분석법에 따라, 여섯 개의 파라미터로 구성되는 어파인 모션 모델을 수학적인 모델로 사용하고, 2D 모션에 의해 구해진 영상의 픽셀 좌표를 독립변수 값으로 하는 회귀식(regression equation)을 구성한다. 그리고 회귀분석법을 이용하여 추출한 어파인 모션 모델의 파라미터를 해석하여 영상 내에서 같은 모션 모델을 가지는 영역을 식별한다.

우선 블록 내에 있는 픽셀 좌표 $\phi^T=[1,x,y]$를 독립변수로 보고, i번째 있는 픽셀의 어파인 모션 모델 파라미터를 $a_{ix}{}^T$와 $a_{iy}{}^T$라 하자. 결국, 어파인 모션 모델을 구하기 위한 회귀 식을 식(2)와 같이 구성할 수 있고, 식(3)으로 블록 B_i의 어파인 모션 모델 파라미터를 구할 수 있다. $u_{2D}(x,y)$와 $v_{2D}(x,y)$는 다단계 웨이블릿 변환 영상으로부터 추정된 2D 모션 값이다.

$$u_{2D}(x,y) = \phi^T a_{ix}$$
$$v_{2D}(x,y) = \phi^T a_{iy} \quad \text{식(2)}$$

$$[a_{iy}\ a_{ix}] = \left[\sum_{B_i}\phi\phi^T\right]^{-1} \cdot \sum_{B_i}(\phi[v_{2D}(x,y)\ u_{2D}(x,y)]) \quad \text{식(3)}$$

또한 식(4)와 같이 회귀 식에서 계산된 예측 값과 측정 값 사이의 오차의 평균값으로 초기 모션 모델로 사용할 것인지의 여부를 판단한다[5].

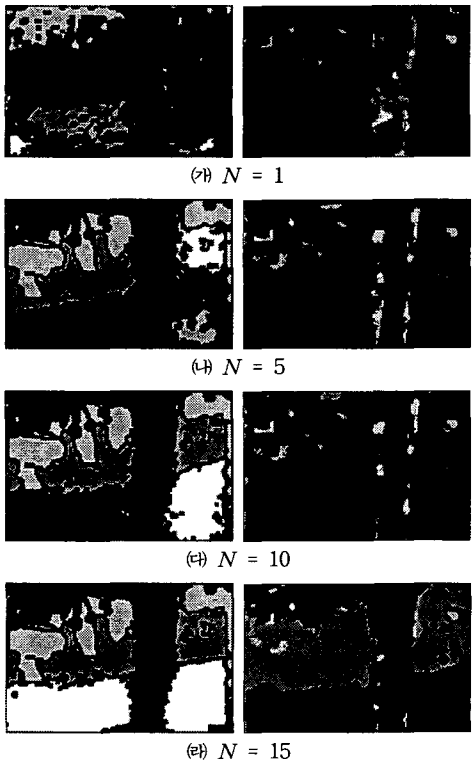
$$\sigma^2 = \frac{1}{N_i}\sum_{(x,y)\in B_i}[(u_{2D}(x,y)-u_a(x,y))^2 + (v_{2D}(x,y)-v_a(x,y))^2] \quad \text{식(4)}$$

N_i는 블록 B_i에 포함된 픽셀의 개수를 나타내고, 임계치(threshold)보다 에러 값 σ^2이 작은 블록을 초기 모션 모델로 설정한다.

초기 모션 모델로 결정된 블록들은 블록 내에 서로 다른 모션을 가지는 물체가 존재하지 않는다는 것을 뜻하기 때문에, 모션 모델링 결과를 신뢰할 수 있다는 것을 의미한다.

3.2 모션 모델 클러스터링 기법

본 논문에서의 모델 클러스터링 과정은 어파인 모션 모델 파라미터 공간에서 k-means 클러스터링 알고리즘[4, 7]으로 모델을 병합함으로써 수행된다. 본 논문에서 제시한 모션 모델 클러스터링 기법을 살펴보면 우선, 초기 모션 모델의 파라미터를 클러스터의 중심으로 하여 모델 병합을 수행하고, 유사한 모델 파라미터를 가지는 영역을 병합 한 후에 모션 모델의 파라미터를 재 계산한다. 이것은 클러스터의 중심을 다시 수정하게 되는 효과를 가져다준다. 그리고 재조정된 클러스터의 중심으로부터의 거리가 가까운 픽셀의 2D 모션을 골라서 해당 클러스터의 레이블을 가지게 한 다음, 픽셀에 할당된 레이블이 더 이상 바뀌지 않을 때까지 이러한 과정을 반복 수행한다.



(가) N = 1

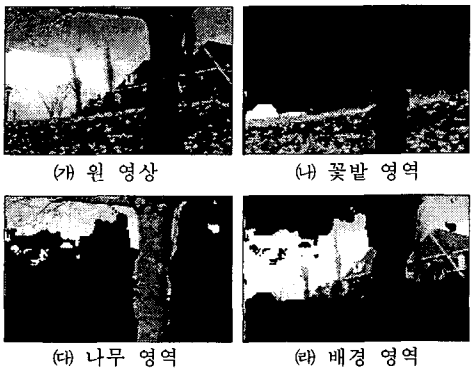
(나) N = 5

(다) N = 10

(라) N = 15

그림 4. 장면 분할 결과 비교

그림 4는 Wang[1]이 제안한 적응적 어파인 클러스터링 알고리즘(왼쪽)과 본 논문에서 제시한 알고리즘(오른쪽)에 의한 장면 분할 결과를 클러스터링 반복 횟수별로 서로 비교한 것이다.



(가) 원 영상 (나) 꽃밭 영역

(다) 나무 영역 (라) 배경 영역

그림 5. 원 영상과 분할된 영상

그림 5는 제시한 알고리즘으로 생성된 분할된 장면들을 보여준다. 입력 영상 프레임을 꽃밭 영역 및 나무 영역, 배경 영역 등으로 구분하여 나타내었다.

Ⅳ. 결론 및 향후과제

본 논문에서는 정확한 2D 모션 모델링을 위해 다단계 웨이블릿 변환 영상으로부터 2D 모션을 계층적으로 추정하는 방법을 제시하였다. 또한 카메라의 움직임에 따른 영상의 2D 모션을 어파인 모션으로 모델링하여 해석하고, 이를 통해 모션 분할 알고리즘을 적용함으로써, 영상 프레임을 같은 평면(planar)상에 존재하는 물체들끼리 분리할 수 있음을 보였다.

특히, Wang[1]이 제안한 알고리즘은 장면을 분할한 후, 영상의 픽셀 값으로 분할된 결과를 조정하는 후처리 과정을 거치는 반면, 본 논문에서는 그러한 별도의 과정 없이 장면을 분할할 수 있었다. 그러나 초기 모션 모델의 결정이 장면 분할 결과에 영향을 미칠 수 있기 때문에 보완해야 할 필요가 있다.

제안한 시스템에 사용된 알고리즘은 멀티미디어 데이터의 조작 및 편집 등에 이용될 수 있으며, 특히, 컴퓨터 그래픽스 분야에서 영상 기반 가상 환경의 자동 생성 및 가상환경 내의 물체 조작을 위한 기초 기술로 사용될 수 있다고 사료되어 진다.

참고문헌

[1] J. Y. A. Wang and E. H. Adelson. "Representing moving images with layers," *IEEE Trans. Image Processing*, vol. 3, pp. 625-638, Sept. 1994.

[2] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms*, Prentice Hall, 1998.

[3] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1999.

[4] A. M. Tekalp, *Digital Video Processing*, Prentice Hall, 1995.

[5] Y. Altunbasak and E. Eren, "Region-based affine motion segmentation using color information," Graphical Models and Image Processing, vol. 60, no. 1, pp. 13-23, Jan. 1998.

[6] G. D. Borshukov, G. Bozdagi, Y. Altunbasak, and A. M. Tekalp, "Motion segmentation by multistage affine classification," *IEEE Trans. Image Processing*, vol. 6, no 11, pp. 1591-1594, Nov. 1997.

[7] R. Klette, K. Schlüns, and A. Koschan, *Computer Vision: Three-Dimensional Data from Images*, Springer, 1998.