

Wavelet 기반의 Watershed를 이용한 효율적인 영상 분할 기법

김종배 김항준

경북대학교 컴퓨터공학과 인공지능연구실

(kjblove, kimhj)@ailab.kun.ac.kr

Efficient Image Segmentation using Wavelet-based Watershed

Jong-Bae Kim^{*} Hang-Joon Kim

Dept. of Computer Engineering, Kyungpook National University

요 약

본 논문은 wavelet 기반의 watershed를 이용한 효율적인 영상 분할을 기법을 제안한다. 영상 분할을 위해 입력 영상을 wavelet transform을 사용하여 low-resolution 영상을 생성한 후 watershed 알고리즘을 이용해 분할하고, 이를 Inverse wavelet transform함으로써 원 영상으로 복원한다. 복원된 영상을 의미 있는 영역들로 분할하기 위해 wavelet 특징값의 유사성을 두 인접한 영역에 비교하여 병합한다. 실험 결과 제안한 방법은 영상의 잡음에 대한 강인함과 영상의 과분할 문제를 해결할 수 있다.

1. 서론

영상 분할은 입력된 영상을 유사한 화소들의 집합인 영역들로 화소들을 구분하는 작업이다. 영상 분할은 비전 기반의 자동 로봇, 상품의 질 검사, 의학에서의 진단과 같은 하위 레벨의 비전 문제 뿐만 아니라, 움직임 분석 및 추적과 같은 연속된 영상 처리, 물체 인식 및 장면 해석과 같은 문제들의 수행 결과에도 고도의 영상 분할 기법에 의존한다. 특히 표준화된 동영상 압축 기법인 MPEG-4/7 영상 압축에 있어서도 의미 있는 객체 분할이 중요한 작업 중에 하나이다. 현재까지 화상 통신과 같은 영상 전송에 있어서 실시간 영상 처리와 함께 효율적인 압축 및 복원 기법에 관련된 연구가 진행되고 있으나, 이러한 작업들을 위해서는 영상에 포함된 잡음을 영향을 최소화하고, 효율적이면서 실시간 처리가 가능한 영상 분할을 기법이 필요하다. 영상 분할을 위한 방법으로 여러 가지 알고리즘이 제안되어 왔다. 영상 분할 방법은 물체와 물체 또는 물체와 배경 사이의 경계를 검출하는 경계 기반 알고리즘 (boundary-based algorithm) [1]과 물체 내의 화소값의 유사성을 이용하는 영역 기반 알고리즘 (region-based algorithm) [2], 수리형태론 기반 알고리즘 (mathematical morphology-based algorithm) [3]을 이용하는 방법으로 구분할 수 있다. 여러 가지 영상 분할 방법들 중에서 수리형태론 기반의 영상 분할은 영상의 크기나 형태등과 같은 특징을 이용하여 기하적인 신호처리 도구로 많이 사용되고 있다. 하지만, 수리형태론 방법은 영상 내에 포함된 잡음에 민감하게 반응하여 영상이 과분할되는 결과를 초래하기도 한다. 수리 형태론 기법중에 watershed 방법을 영상 분할에 적용하기 위해서는 영상에 포함된 잡음을 제거함과 동시에 영상 전송 및 압축과 같은 작업에서 효율적인 영상 처리 기술이 필요하다.

본 논문에서는 Wavelet Transform (WT)을 이용하여 Pyramid 영상을 생성하고, 생성된 Pyramid의 low-resolution 영상을 watershed 알고리즘을 사용하여 영상을

분할한다. 그리고 분할된 영상을 inverse WT하여 원 영상으로 복원한다. 복원된 분할 영상에서 각 영역의 wavelet feature의 유사성을 비교하여 영역을 병합함으로써 영상에서 있는 의미 있는 물체 영역으로 분할하는 방법을 제안한다. 잡음에 강한 영상 분할을 위해 WT을 이용하여 low-resolution 영상을 생성함으로써 영상의 noise filtering 기능을 제공할 뿐만 아니라 영상 분할에서 영역의 과분할 (over segmentation) 문제를 극복할 수 있다. 그림 1은 제안한 방법의 알고리즘을 나타낸 그림이다.

2. 영상 분할

본 단계에서는 이미지 sequence $I(k)$ 를 입력하여 유사한 특징을 가진 영역들로 분할한 영상 $S_0(k)$ 을 출력한다. 구성된 단계는 원 영상 $I_0(k)$ 를 WT을 사용하여 Pyramid 영상을 생성한다. 그리고 Pyramid 영상의 low-resolution 영상 $L_L(k)$ 을 Watershed 알고리즘에 의해 유사한 영역들로 영상을 분할 하고 $S_L(k)$, Inverse WT를 이용하여 분할된 원 영상 $S_0(k)$ 으로 복원한다.

2.1 Wavelet Transform을 이용한 Pyramid 영상 생성

WT은 최근 들어 신호 처리나 패턴 인식 분야에서 가장 많이 쓰이고 있는 변환 방법중의 하나이다. WT은 영상을 주파수 대역별로 분할하고, 분할된 저대역(lowpass) 부밴드(subband)는 원 영상이 가지는 에너지를 대부분을 소유하고 있으며, 고대역(highpass) 부밴드는 인간 시각에 민감한 영상의 에지 정보를 보유하고 있다. 특히 WT은 노이즈 감소에도 적용하고, 또한 low-resolution 영상 생성에도 이용되어지고 있다 [4]. 제안한 방법은 Haar WT을 이용하여 Pyramid 영상을 생성한다. WT은 영상에 포함되어 있는 노이즈의 영향을 최소화 시키는 동시에 영역 분할의 과분할되는 결과를 줄이기 위해 영상의 차원을 줄이게 사용한다.

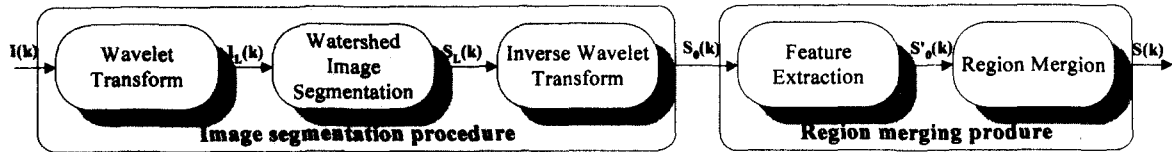


그림 1. 제안한 영상 분할 방법의 Block diagram. $I(k)$ 는 k 시간일 때의 이미지 sequence, $I_L(k)$ 는 L 번째 low-resolution 이미지, $S_L(k)$ 는 각 processing 단계의 분할된 이미지

Pyramid 영상 생성의 첫 단계는 입력 영상 $I(k)$ 에서 WT를 이용하여 Pyramid 영상을 생성하는 것이다. 여기서 k 는 연속된 이미지에서 시간을 의미한다. k 시간일 때, 원 영상 $I_0(k)$, 0는 full-resolution image ($I(k) = I_0(k)$), 각 low-resolution image ($I_1(k), I_2(k), I_3(k), \dots, I_L(k)$)들이 생성된다. 그림 2는 WT를 사용한 Pyramid 영상을 생성하는 hierarchical 구조를 나타낸다.

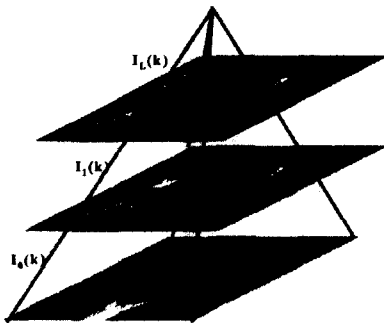


그림 2. Wavelet transform을 이용한 Pyramid 영상 생성. $I_0(k)$ 는 full-resolution 영상, $I_L(k)$ 는 low-resolution 영상

2.2 Watershed를 이용한 영상 분할

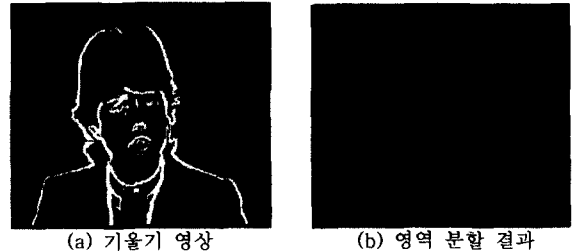
Watershed 알고리즘은 지형학 분야에서 연구되어 온 것으로서, 단순화된 영상에 대한 경사 영상 화소의 밝기 값을 고도 정보로 생각함으로써 영상 처리 분야에 응용되기 시작했다. 이는 담수 지역 또는 마커 영역을 구분하는 분수령을 찾음으로써 각각의 국부 최소치에 연관된 영역을 분할해 내는 알고리즘이다. 그 동안 지금까지는 분수령을 계산하기 위한 다양한 알고리즘들이 제안되었으나, 그 중 immersion simulation에 의한 방법이 가장 효과적인 기법으로 알려져 있으며, 이것은 계층적 큐(hierarchical queue)로 구현이 가능하다 [3]. Immersion simulation을 이용한 영상 분할 과정은 다음과 같다. 우선 가장 고도로부터 시작하여 담수 지역에 물을 차례로 채워나간다. 물이 다른 국부 최저 위치에 이르면 이 두 지역에 동시에 물이 채워져 나간다. 서로 다른 지역의 물이 채워지다가 만나게 되면, 만나는 경계선에 가상의 댐을 쌓는다. 이 과정이 모두 끝나면 물의 위치를 절대 최고치의 높이에 이르게 되는데, 이 때 물로 채워진 각 국부 최소치의 영역이 담수 지역이고, 담수 지역들을 둘러싸는 댐의 선이 분수령이 된다.

그림 3은 watershed 알고리즘을 이용한 영역 분할 그림이다. 그림 (가)는 Claire 136번 영상의 기울기 영상이고, (나)는 영역이 분할된 결과이다. (다)는 기울기 영상을 3차원으로 표현한 그림이다.

2.3 Inverse WT에 의한 분할된 영상의 복원

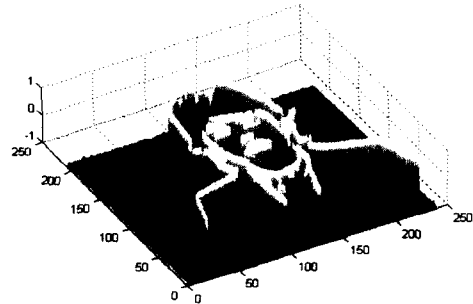
영상 분할 단계에 의해 분할된 low-resolution 영상 $S_L(k)$ 을 Inverse WT에 의해 분할된 원 영상 $S_0(k)$ 으로 복원한다. 기존의 low-resolution 영상의 복원 방법에는 직접

projection에 의한 방법 [5], relaxation 방법 [6]이 주로 이용되고 있다. 그러나, 직접 projection 방법은 영상의 없는 정보를 다시금 복원해야 하는 문제로 인해 영상의 질을 저하시킬 수 있으며 relaxation 방법은 복원 되는 계산 시간이 많이 소모되는 문제점이 있다. 하지만 제안한 방법에서 영상 복원을 위해 inverse WT를 적용함으로써 원 영상의 품질 저하를 방지하면서 영상 복원에 걸리는 시간을 relaxation 방법보다 감소시킬 수 있다.



(a) 기울기 영상

(b) 영역 분할 결과



(c) 3차원 영상

그림 3. watershed 알고리즘의 영상 분할

3. 분할된 영역의 특징 추출 및 영역 병합

영상을 의미 있는 영역들로 분할하기 위해서 watershed를 의해 분할된 영역을 다시금 영역 병합 과정을 거쳐야 한다. 이를 위해 분할된 인접한 영역의 wavelet 특징값 유사도를 비교하여 병합한다. 분할된 영역의 wavelet 특징 값은 영상 분할 단계에 의해 분할된 각 영역에 대응하는 L -level wavelet decomposition 영상의 저주파 영역에 평균(M)과 2, 3차 모멘트 (μ_2, μ_3)를 사용한다. 특징 값들은 식(1, 2, 3)과 같이 계산한다. W_L 는 분할된 영역에 대응하는 L -level WT 저주파 영역이다. m, n 는 분할된 각 영역의 대응하는 저주파 영역의 픽셀들이다. 두 인접한 두 영역의 wavelet 특징 값의 합이 임계치 이상이면 분할된 두 영역을 병합한다.

$$M = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} W_L(i, j) \quad (1)$$

$$\mu_2 = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (W_L(i, j) - M)^2 \quad (2)$$

$$\mu_3 = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (W_L(i, j) - M)^3 \quad (3)$$

4. 실험 결과

제안한 방법은 MPEG-4/7의 표준화 과정에서 사용된 테스트 영상 중, 'Table Tennis', 'Claire', 'Hall monitor', 'Akiyo', 영상을 사용하였다. 각 실험 영상은 176×144 크기의 QCIF format 영상을 사용하였고, 실험 환경은 Pentium IV 1.4Ghz PC에서 Matlab 5.3으로 실험하였다. 한 프레임 당 평균 처리 시간은 0.23초였다. Watershed 알고리즘의 초기 담수화가 이루어지는 마커 영역은 기울기가 0이고 크기가 10 이상인 영역으로 지정하였다. 그림 4는 원 영상을 watershed 알고리즘에 의해 분할 결과 영상과 1-level WT후 저주파 영역의 low-resolution 영상을 분할한 결과이다. 성능 차이를 비교하기 위해서 분할된 영역의 갯수와 분할된 영역을 평균값으로 대체한 복원 영상의 PSNR를 측정하였다. 표 1은 원 영상과 low-resolution 영상에서 watershed를 이용하여 분할된 영역의 갯수와 PSNR를 나타낸 결과이다. 실험에서 low-resolution 영상에서 분할한 결과 영상이 원 영상 보다 과분할 문제를 감소 시킬 뿐만 아니라 복원된 화질 차이도 원 영상을 복원한 결과에 비해 차이 거의 없음을 알 수 있다. 표 2는 제안한 방법의 노이즈 민감성 측정 결과이다. 실험 결과 제안한 방법이 원 영상 처리하는 것보다 노이즈에 강함을 알 수 있다.

표. 1 제안한 방법에 영상 분할 성능 측정 (분할된 영역의 갯수 / PSNR (dB))

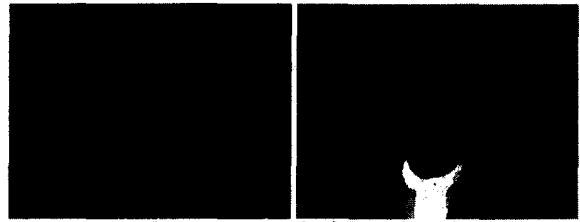
Test 영상 (Maker size)	원 영상	1-level WT	2-level WT
Tennis 2번(20)	790 / 27.31	63 / 27.03	27 / 27.75
H/M 58번(17)	285 / 26.92	130 / 27.66	66 / 26.12
Claire 136번(15)	46 / 27.18	38 / 27.67	31 / 27.23
Akiyo 2번(10)	492 / 28.63	132 / 28.91	65 / 28.73

표. 2 노이즈 민감성 실험 (원 영상의 PSNR (dB) / 1-level WT의 PSNR (dB))

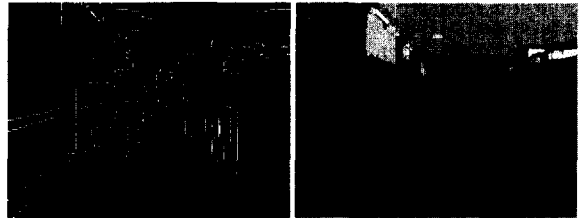
노이즈 (%)	0	10	20
H/M 58번	26.92 / 27.66	24.72 / 25.30	22.12 / 23.89
Claire136번	27.18 / 27.67	23.42 / 25.02	21.37 / 24.72
Akiyo 2번	28.63 / 29.91	24.53 / 25.73	22.25 / 23.17



(가) 'Table Tennis' 2번 영상 분할 결과



(나) 'Claire' 136번 영상 분할 결과



(다) 'Hall Monitor' 58번 영상 분할 결과



(라) 'Akiyo' 2번 영상 분할 결과

그림 4. 제안한 영상 분할 결과. 1-level WT의 저주파 영역을 watershed로 분할한 결과(왼쪽), 영상 복원(오른쪽)

5. 결론

본 논문에서는 실제 비디오 영상에서 의미 있는 객체를 효율적으로 분할하는 방법을 제안하였다. 영상 분할을 위해 원 영상에 wavelet transform을 이용하여 low-resolution 영상을 생성한 후 watershed 알고리즘을 이용해 분할함으로써 영상의 잡음에 대한 강인함과 영상의 과분할 문제를 극복할 수 있다.

6. 참조 논문

[1] D. H. Ballard and C. M. Brown, Computer Vision, Prentice-Hall, Inc., EngleWood Cliffs. Nj. 1992.
 [2] N. R. Pal and S. K. Pal, 'A reviews On Image Segmentation Techniques', Pattern Recognition., Vol. 26, No. 9, pp. 1277-1294, 1983.
 [3] L. Vincent and P. Soille, 'Watersheds in Digital Spaces: An Efficient Algorithms Based on Immersion Simulations', IEEE Trans. PAMI, Vol. 13, No. 6, pp. 538-598, 1991.
 [4] M. Lang, H. Guo, J.E. Odegard, C.S. Wells, ' Noise reduction using an undecimated discrete wavelet transform' IEEE Signal Processing Letters, Vol. 3 No. 1, pp. 10 - 12, 1996.
 [5] Gonzalez, 'Digital Image Processing', Addison-Wesley
 [6] L. Salgado, N. Garcia, L.M. Menendez and E. Rendon, 'Efficient Image Segmentation for Region-Based Motion Estimation and Compensation', IEEE Trans. Circuits and Syst. Video Tech. Vol. 10, No. 7, pp. 1029-1039, 2000.