

Hausdorff 측도를 이용한 영역기반 움직임 벡터 추정

임봉일, 최윤식

연세대학교 전기공학과

Region-based Motion Vector Estimation Using Hausdorff Measure

Bong-il Yim, Yoon-sik Choe

Department of Electrical Eng. Yonsei University

요약문

최근에는 영역(혹은 객체)를 이용하여 비디오 시퀀스를 표현하거나 부호화하는 기법들이 많이 연구되고 있다. 이러한 부호화 기법에서는 형태정보를 효율적으로 이용하는 것이 중요함에도 불구하고, 현재 사용되고 있는 대부분의 기법에서는 기존의 블록 기반 부호화 알고리즘에서처럼 오직 PSNR 만을 고려하여 움직임 벡터를 추정하고 있다. 따라서, 형태 정보를 다루는 효율적 움직임 추정 알고리즘이 필요하다. 본 논문에서는 각 영역의 경계(contour)를 잘 피팅(fitting)시키는 움직임 추정 방법을 생각해 본다. 이를 위하여 PSNR과 영역의 모양을 함께 고려하는 비용함수를 제안하고 이를 이용한 움직임 벡터 추정을 고려해 본다.

1. 서론

현재 사용되는 MPEG2에 의해 압축된 데이터는 압축률이 높은 반면 PSTN이나 이동통신망, 인터넷망을 이용한 실시간 전송에는 부족함이 있으며, 사용자와의 상호작용을 전혀 고려하고 있지 않다. 따라서, 새로운 표준이 필요하였으며 그 결과가 MPEG4/7이다. MPEG4에서는 4.8~64kbps 범위의 초저속 전송망에서의 전송을 위한 높은 압축율을 얻기 위해서 만이 아니라 MPEG4에서 제공하고자 하는 다양한 기능(내용 기반 상호작용성, 범용적인 접근 용이성)을 제공하기 위하여 VOP 즉, 임의의 모양을 갖는 영역 단위로 부호화가 이루어진다.

현재 사용되는 영역기반 움직임 벡터 추정 방법을 살펴보면, 움직임 벡터의 추정 단위가 임의의 모양을 갖는 객체라는 것이 다들 뿐 기존의 블록 기반 방법과 같은 방법을 사용한다. 이 때 에러의 많은 부분은 경사도(gradient)값이 큰 에지영역을 중심으로 고주파 영역에서 나타나게 된다. 따라서 에지영역을 잘 근사할 수 있다면 에러의 분포 범위를 줄일 수 있고 전체 발생 비트율에도

영향을 주게된다. 그러나 기존의 방법들은 PSNR이 최대가 되도록 움직임 벡터를 찾는 반면 이러한 에지영역을 어떻게 하면 잘 근사할 수 있는가에 대한 고려는 전혀하지 않고 있다. 그러므로 움직임이 큰 동영상의 경우, 움직임보상 예측한 영상을 살펴보면 에지부분이 거칠게 표현되는 것을 볼 수 있으며 에러의 분산값이 크다는 것을 알 수 있다. 따라서, 에지 모양을 잘 근사할 수 있는 측도를 생각해 보고 이것을 이용하여 움직임 벡터를 찾는 방법을 생각해 보는 것이 필요하다. 에지 모양의 최적 근사를 고려한 방법으로 움직임 보상을 하는 경우 평균 에러값은 기존의 방법보다 커질 수 있고, 움직임보상 예측한 영상의 PSNR은 상대적으로 낮아질 수 있다. 하지만 에러의 분산값은 줄어 들 수 있기 때문에 결과적으로 전체 발생 비트율에 영향을 줄 수 있다. 본 논문에서는 윤곽선 모양을 잘 근사할 수 있는 측도로서 Hausdorff distance를 생각해 보고, PSNR과 윤곽선 모양의 효율적인 근사를 함께 고려하는 새로운 비용함수를 제시한다. 그리고, 이를 이용하여 움직임 벡터를 추정한 후 에러의 분산값이 작아지는 것을 살펴보고자 한다.

2. Hausdorff measure

Hausdorff distance는 두 집합 A, B가 서로 얼마나 가까이 근접해 있는가를 측정하는 척도이다. 따라서 두 객체가 겹쳐져 있고 객체의 경계(boundary)를 나타내는 집합을 고려하는 경우에는, 겹쳐진 객체의 모양이 서로 얼마나 닮았는가를 결정하는 수단으로 사용할 수 있다.

Hausdorff distance의 정의는 다음과 같다. 우선 X에서 A까지의 거리 $d_A: X \rightarrow R^1$ 를 다음과 같이 정의하자.

$$d_A(x) = d(x, A) = \min \{ d(x, a) : a \in A \} \quad (1)$$

이 함수 d_A 는 임의의 집합 B 위에서 최대값을 갖는다.

이 최대값을 다음과 같이 정의한다.

$$d(B \rightarrow A) = \max_{b \in B} \{d(b, A)\} = \max_{b \in B} \min_{a \in A} \|b - a\| \quad (2)$$

마찬가지로 집합 A위에서 d_B 의 최대값을 $d(A \rightarrow B)$ 로 정의한다. 이제, 이 $d(B \rightarrow A)$ 와 $d(A \rightarrow B)$ 를 이용하여 Hausdorff distance를 다음과 같이 정의할 수 있다[2].

$$h(A, B) = \max\{d(B \rightarrow A), d(A \rightarrow B)\} \quad (3)$$

여기서, $d(B \rightarrow A)$ 는 B에서 A로의 directed Hausdorff distance라고 한다. 이 Hausdorff distance는 집합 A에 속하는 점들이 집합 B로부터, 혹은 역으로, 최대 얼마나 떨어져 있는가 하는 거리를 구함으로써 두 집합의 불일치(mismatch) 정도를 나타내고 있다.

상관도(correlation)나 템플릿 매칭(template matching) 등은 두 객체 모양의 서로 다른 정도를 측정하기 위하여 패턴인식 분야에서 많이 사용하는 기법이다. 그러나 이런 기법들은 측도(metric)가 되기 위한 조건을 만족시키지 못하며, 특징점을 찾아낼 때(feature extraction) 생길 수 있는 특징점 위치의 작은 오차에 민감하다. 반면 Hausdorff distance $h(A, B)$ 는 모든 닫혀있고 유계인 집합(closed bounded set) 위에서 측도(metric)이다[3]. 따라서 Hausdorff distance는 일반 영상을 표현하는 집합 위에서 측도임을 알 수 있다.

Hausdorff distance를 이용하여 두 집합의 모양의 닮음 정도를 나타내기 위해서는, 우선 두 집합의 경계가 이진 영상(binary image)로 표현되어야한다. 그런 다음 하나의 집합을 움직이면서 두 집합의 가능한 모든 위치에 대하여 거리를 계산하고 이 값의 최소값을 찾는다. 예를들어 평행이동(translation) t 만이 고려되었을 때는 Hausdorff distance의 최소값을 다음과 같이 표현할 수 있다.

$$M_T(A, B) = \min_t h(A, B \oplus t) \quad (4)$$

여기서 $B \oplus t = \{b + t | b \in B\}$ 를 의미한다.

3. 영역분할 기반 움직임 벡터 추정 방법

이 실험에서는 에지를 찾고 이 에지로부터 영역을 나누는 방법을 사용한다. 먼저 Canny 에지 검출기를 이용하여 에지를 구한다. 이렇게 구한 에지들은 서로 연결되지 않은 끊어진 에지들이 나타난다. 이렇게 끊어진 에지점들

을 중심으로 8-neighbor에서 경사도를 비교하여 이중 경사도 값이 가장 큰 점들을 연결해 나간다. 경사도 값을 비교할 때 일정한 문턱치 T_2 이상이 되는 점들만을 비교한다. 너무 작은 경사도 값을 나타내는 경우는 영역의 경계를 나타내는 부분이라고 볼 수 없기 때문이다.

이렇게 찾은 영역의 경계(contour)는 두가지로 분류된다. 하나는 닫힌 경계(closed contour)이고 다른 하나는 열린 경계(open contour)이다[4]. 이중 닫혀진 경계의 내부만을 영역으로 선택한다. 열린 경계들은 고르지 못한 텍스처 부분이거나 잡음 때문에 발생하는 것으로 간주한다.

대다수의 영역 기반 부호화 방법에서는, 원영상과 예측 영상 사이의 MSE(Mean Square Error)를 비용함수로 이용하여 움직임 벡터를 찾는다. 그러나 이 때 대부분의 예측 오차는 평탄 영역보다는 윤곽선 영역을 중심으로 넓은 영역에 걸쳐 나타난다는 사실을 생각해 보면, 단순히 화소의 밝기값만을 고려하기 보다는 영역의 윤곽선을 잘 근사할 수 있는 움직임 벡터 추정 방법을 생각해 보는 것이 필요하다는 것을 알 수 있다. 본 논문에서는 영역의 모양의 유사성을 측정하기 위한 측도로서 Hausdorff distance를 이용하는 새로운 비용함수를 생각해 봄으로써 영역의 윤곽선을 고려하는 움직임 추정 보상 방법에 대해 고찰해 보도록 한다. 우선 다음과 같은 비용함수를 제시한다.

$$F_{\cos t} = \frac{1}{N} \sum_{(x,y) \in R_i} [I(x,y,T) - I(x-dx,y-dy,T-1)]^2 + \lambda \cdot h(M(B_{R_i}(T-1)), B_{\cup R_i}(T)) \quad (5)$$

여기서 N은 영역 R_i 내에 속한 화소의 총 개수를 나타내며, $I(x,y,T)$ 는 T프레임에서 화소점 (x,y) 에서의 밝기값을, (dx,dy) 는 T-1 프레임과 T 프레임 사이에서 화소점 (x,y) 의 움직임을 나타낸다. 또한 λ 는 DFD(Displaced Frame Difference)와 Hausdorff distance 사이의 가중치를 나타내며, $B_{R_i}(T-1)$ 는 T-1프레임에서 영역 R_i 의 경계(Contour)를 $B_{\cup R_i}(T)$ 는 T 프레임에서 전체 영역의 경계를 나타낸다. 그리고, $M(B_{R_i}(T-1))$ 는 $B_{R_i}(T-1)$ 의 움직임을 고려했을 경우의 윤곽선을 나타낸다.

이 비용함수를 최소화하는 문제에서는, 일정영역을 탐색하여 최소값을 찾는다. 우선 평행이동(translation)만을 고려하여 최소값을 찾은 다음 이 위치를 중심으로 회전(rotation)을 고려한다. 이때의 회전은 각 영역의 중심점을 기준으로한 회전을 의미하며, 회전각도는 이 중심점에서 가장 먼 거리에 있는 영역 내부의 점이 회전에 의해 1화

소 정도 움직이도록 선택한다. 즉 $\Delta\theta = \arctan(1/r_{\max})$ 를 만족하도록 회전각을 선택한다. 실험에서는 각 영역 내부에 속한 화소로부터 중심점까지의 거리를 직접 계산하지 않고 알파평면을 이용하여 중심점에서부터 알파평면을 이루는 4개의 변까지의 거리중에서 가장 큰 값을 r_{\max} 로 사용한다.

4. 실험결과

본 논문에서는 310×280 Claire 영상을 이용하여 실험을 하였다. 우선 Canny 에지 검출기를 사용하여 에지를 찾는다. 그런 다음 끊어진 점들을 연결한다. <그림4-3>에서 얼굴과 팔 부분에서 열린 경계가 나타나고 나머지 대부분의 경우에는 닫힌 경계가 나타나는 것을 볼 수 있다. 이 경우 닫힌 경계의 내부만을 영역으로 분류해 낸다. 이런 방법으로 구한 영역의 총 개수는 147개였다. 이렇게 구한 영역에 대하여 라벨링(labeling)을 한다.

다음에는 이렇게 얻은 그래프를 바탕으로 영역을 합치는 과정을 거친다. 첫째, 영역에 속한 전체 화소의 개수가 기준 문턱치(threshold)보다 작은 영역은 인접한 영역들 중에서 평균 그레이값을 비교하여 가장 가까운 값을 나타내는 영역으로 합친다. 이 과정을 거친 결과, 총 영역의 개수는 60개로 줄어들었다. 둘째, 인접한 영역들 중에서 같은 움직임 벡터를 갖는 영역들을 합쳐 나간다. <그림4-5>은 최종적으로 얻어진 영역의 평균값을 나타내고 있다. 이렇게 해서 최종적으로 구한 영역들은 임의의 위치와 형태를 갖게 되므로 <그림4-6>에서 보는 것처럼 알파평면을 이용하여 각 영역을 나타내며, 이를 단위로 하여 움직임 벡터를 찾는다.

다음 실험에서는 Hausdorff distance를 이용해서 영역의 경계를 추적해 갈 때 얼마나 강인성(robustness)을 보이는가를 살펴보고자 한다. 실험은 원영상이 수축되었을 때의 영상, 그리고 원영상이 잡음에 의해 약간 변형되고 회전하였을 때의 영상에 대하여 Hausdorff distance가 얼마나 강인하게 물체의 경계를 추적할 수 있는가를 살펴본다. 이 때 실험 결과가 <그림 4-10~4-11>에 나타나 있다. 이 실험 결과를 살펴보면 2가지 실험에서 모두 만족할 만한 결과를 나타내고 있다는 것을 알 수 있다.

다음 실험에서는 PSNR만을 고려한 방법으로 움직임 보상 예측한 영상과 제안한 비용함수를 사용하여 움직임 보상 예측한 영상을 비교해 본다. Claire 시퀀스 제 2프레임으로부터 제 20프레임을 움직임 보상에 예측하면 각각 <그림4-14>, <그림4-15>와 같다. 영역기반 움직임 추정에서

는 블록효과가 나타나지 않는 대신에 영역별로 움직임이 다름에 따라 드러난 영역과 겹쳐진 영역이 생기게 된다. 그림에서 얼굴과 머리 부분에 겹게 갈라져 보이는 부분이 드러난 영역을 나타낸다. 이 실험에서는 이들 영역에 대한 보정을 목표로 하는 occlusion processing은 고려하지 않는다. 제안한 비용함수를 사용하여 움직임 보상을 할 때, λ 값은 실험적으로 결정하였으며, 이 실험에서는 λ 값으로 500을 사용하였다. <그림4-14>와 <그림4-15>에서 살펴보면 제안한 방법으로 움직임 보상을 한 경우가 얼굴과 입 영역의 윤곽선이 잘 표현되고 있음을 알 수 있다.

실험에서는 Hausdorff distance를 구할 때, 비교되는 두 영역의 경계를 나타내는 각각의 집합에서 서로 가장 잘 일치하는 70%만을 선택하여 사용하였다. 각 방법에서 에러의 분산값을 비교해 보면 <표4-1>과 같다.

5. 결론

객체에 기반한 방법으로 부호화를 수행하는 제 2세대 부호화 방법이 기존의 제 1세대 부호화 기법들과 가장 크게 다른점은, 움직임 정보와 텍스처 정보 외에 영역의 형태에 대한 정보가 알파평면과 같은 방법으로 표현되어야 한다는 것이다. 그러나 현재 사용되는 움직임 추정 보상 방법은 기존의 블록 기반 방법을 그대로 확장하여 PSNR만을 고려하고 있다.

본 논문에서는 Hausdorff distance를 영역의 모양의 유사도를 측정하는 측도로 채택하여, 에러의 에너지와 형태의 유사성을 함께 고려하는 비용함수를 생각해 보고, 이를 이용하여 영역의 움직임을 추정하는 방법을 생각해 보았다. 또한 이 경우 영역의 에지 부분을 잘 근사함에 따라 원영상과 움직임 보상 예측한 영상 간의 에러의 분산값이 줄어든다는 것을 살펴보았다.

참고문헌

- [1] Heinz-Otto Peitgen, Hartmut Jurgens, Dietmar Saupe, "Chaos and Fractals : New Frontiers of Science", Springer-Verlag, 1992
- [2] Allan J. Sieradski, "An Introduction to Topology and Homotopy", PWS-KENT publishing Company, 1992
- [3] Daniel P. Huttenlocher, Gregory A. Klanderman, and William J. Rucklidge, "Comparing Images Using the Hausdorff Distance" IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 15, No. 9, September 1993, pp. 850-863

- [4] Hui Li, B. S. Manjunath and Sanjit K. Mitra, "A Contour-Based Approach to Multisensor Image Registration ", IEEE Trans. on Image Processing, vol. 4, No. 3, March 1995, pp. 320-334
- [5] L.Wu, J. Benois-Pineau, Ph. Delagnes, D. Barba, "Spatio-temporal segmentation of image sequences for object-oriented low bit-rate image coding", Signal processing: Image Communication 8 (1996), pp. 513-543



그림4-1. Claire frame 2



그림 4-2. 에지영상

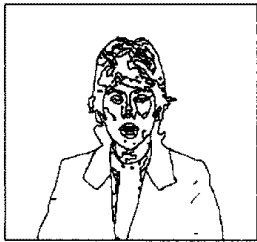


그림 4-3. 에지를 연결한 그림



그림 4-4. 영역 123



그림 4-5. 최종 영역의 평균값

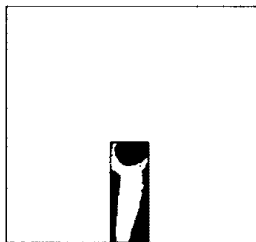


그림 4-6. 영역 123의 알파 평면

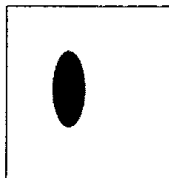


그림 4-7. 원영상

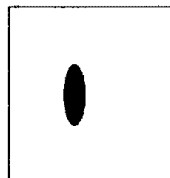


그림4-8. 실험영상1

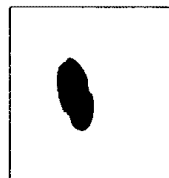


그림4-9. 실험영상2

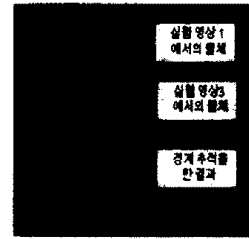


그림4-10. 경계추적결과1



그림4-11. 경계추적결과2

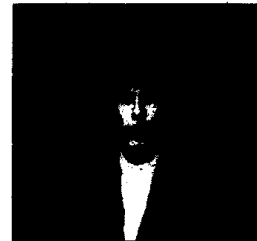


그림 4-12. frame 2

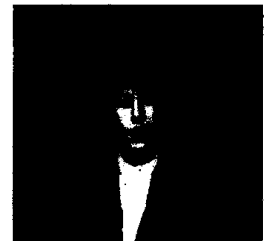


그림 4-13. frame 20



그림 4-14. PSNR만을 고려한 방법

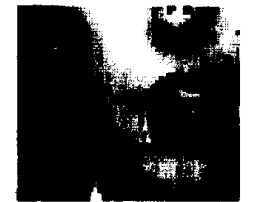


그림 4-15 제안한 비용함수를 이용한 방법

표 4-1. variance 비교

	PSNR 만을 고려한 방법	제안한 비용함수를 이용한 방법
variance	338.68	301.81