투영된 모션과 히스토그램 인터섹션 기법을 이용한 강건한 물체추적

이봉석°, 문영식
한양대학교 컴퓨터공학과

Robust object tracking using projected motion and histogram intersection

Bong Seok Lee and Young Shik Moon
Department of Computer Science & Engineering
Hanyang University, Ansan, Kyonggi Do, 425-791, Korea
bslee@cse.hanyang.ac.kr, ysmoon@cse.hanyang.ac.kr

요약

본 논문에서는 투영된 모션과 히스토그램 인터섹션을 이용한 노이즈에 강건한 물체추적 방법을 제안한다. 기존의 방법은 템플릿 매칭, 물체의 경계선 재 검출, 물체의 움직임 정보 등을 사용하여 물체추적을 하였으나, 템플릿 매칭의 경우 많은 계산 시간을 요구하며 경계선을 재 검출하는 경우 윤곽선이 잘못 설정되는 경우가 있고 물체의 움직임 정보를 사용하는 경우에는 움직이는 카메라에서 움직이는 물체만을 추적하기가 쉽지 않은 단점이 있다. 본 논문에서는 투영된 모션과 질의 영상의 템플릿 마스크를 사용하여 물체의 이동, 회전과 스케일을 고려한 노이즈에 강건한 물체추적 기법을 제안한다. 질의영상은 영상분할 후 영역선택을 통하여 구성하고 물체의 인식은 색상을 이용한 히스토그램 인터섹션 기법을 사용한다. 물체의 이동은 가로 및 세로의 밝기 값을 1차원 신호로 투영하여 개략적인 움직임을 감지하고 이동에 대한 에러를 보정하며 회전과 스케일의 변화는 질의 영상의 템플릿 마스크를 이동하여 회전과 스케일에 맞게 변경하여 감지한다.

1. 서 론

컴퓨터의 대중화, 통신기술의 향상과 영상기기의 급속한 발달로 우리 주변에서는 많은 양의 멀티미디어 자료들을 접하게 되었다. 그 결과로 현재는 인터넷을 통해 동영상을 실시간으로 전달받아 시청 할 수 있다. 이로 인해 동영상을 통해 많은 정보를 전달 하고자 하는 방법들이 개발되고 있으며 본 논문에서도 전자상거래를 위한 동영상에서 상품정보를 제공하는 것을 목표로 한다. 모든 영상 및 비디오는 신호처리의 기본적인 전처리로서 의미 있는 객체의 분할과 객체의 시간적인 움직임 추적이 요구된다. 압축효율의 향상과 다양한 기능을 가능하게 하기 위해, 객체기반 영상압축 기법이 MPEG-4의 국제 표준으로 확정되었다 [1]. MPEG-4는 객체를 단위로 비디오 신호를 압축하지만 객체의 분할은 표준에 포함되어

있지 않고 미래의 기술발전에 의존하고 있다. 현재의 기술로 의미 있는 객체의 자동적인 분할은 많은 어려움이 있다. 그러나 밝기 값, 색상정보, 텍스처 등의 동질성을 갖는 영역의 분할은 가능하다. 본 논문에서는 인간의 개입을 최소화시켜 객체를 추출하는 방법과 빠른 시간에 객체의 움직임을 추적하여 회전, 스케일에 적응적이고 노이즈에 강건한 알고리즘을 제안한다. 본 논문은 크게 영상분할, 물체추적, 객체인식의 세 가지로 나누어진다. 영상분할에 관한 연구는 불연속성, 에지(edge) 연결과 윤곽선, Thresholding, 영역기반, 움직임 등을 이용한 분할기법으로 나눌 수 있다[2]. 본 논문에서는 수리형태학적 연산자를 이용한 영상분할 기법을 이용한다. 물체추적에 관한 연구는 템플릿 매칭 기법[2], Affine 변환 기법[3], 블록정합 기법[4], 투영된 모션 기법[5] 등으로 나눌 수 있다. 템플릿 매칭은 이동, 회전, 스케일 된 템플릿으로 재구성하여 비교하면 어느 정도 나은 결과를 얻을 수 있지만 모든 영역을 탐색하기 때문에 계산 량이 매우 많다는 단점이 있다. Affine 변환을 통한 움직임 추정 기법은 비교적 정확한 결과를 가져올 수 있지만 객체 내의 각 영역에 대하여 움직임 추정에 많은 연산 및 시간을 요구하므로 본 논문에서와 같이 동영상에서 물체 추적을 요구하는 경우에는 적합하지 않다. 블록 정합 기법은 전역탐색 알고리즘과 고속탐색 알고리즘이 있다. 전역탐색 알고리즘은 움직임 추정기법들 중에서 최적의 움직임 벡터를 찾는 장점이 있는 반면에 탐색영역에 포함되는 모든 점들을 탐색하기 때문에 계산 량이 너무 많다는 단점이 있다. 고속탐색 알고리즘은 국부적인 탐색으로 이루어져 있으며 움직임이 거의 없는 영상의 경우 문제가 없지만 움직임이 많은 경우 첫 번째 탐색이 잘못 되었을 경우 local optima에 빠질 수 있는 단점이 존재한다. 투영된 모션 기법은 가로 및 세로의 밝기 값을 1차원 신호로 투영하여 개략적인 움직임을 감지한다. 이 기법의 경우 상당히 빠른 시간에 물체의 움직임 위치를 찾아 낼 수 있으며 약간의 에러는 있지만 템플릿 마스크를 적용하여 히스토그램을 비교하면 에러를 보정 할 수 있다. 객체인식의 경우 컬러특징 값을 이용한 방법, 모양 특징 값을 이용한 방법, 컬러와 텍스처 특징 값을 혼용하는 방법 등이 있다[6][7][8]. 본 논문에서는 영상 전체 면적으로 정규화 한 확률 값을 이용한 변형된 색상 히스토그램 인터섹션 기법을 사용한다.

2. 제안하는 알고리즘

제안하는 알고리즘은 객체추출, 물체추적, 객체인식 순으로 구성된다. 그림 1은 전체적인 순서도를 보여준다.

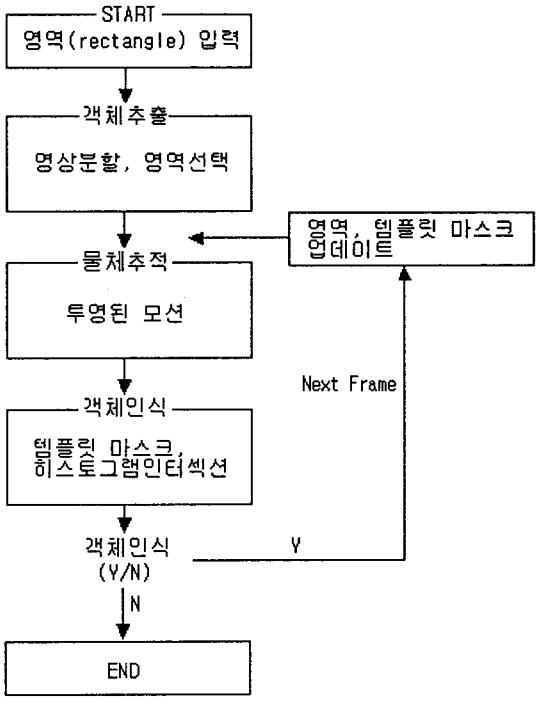

그림 1. 제안된 방법의 전체 순서도

2.1 객체추출

객체추출은 인간이 최소의 개입으로 빠른 시간에 에러 발생의 소지가 없고 비디오의 종류에 관계없이 객체를 추출 할 수 있는 영상분할 및 병합 기법을 사용한다. 영상분할은 비교적 속도와 성능이 좋은 Demin Wang이 제안한 기울기와 워터쉐드(Watershed)를 이용한

기법을 사용한다[9]. 그림 2는 영상분할 한 영상을, 그림 3은 분할된 영상으로부터 마우스를 이용하여 원하는 객체를 추출하는 것을 보여준다.

그림 2. 영상분할 한 영상

그림 3. 객체추출

2.2 투영된 모션 기법을 사용한 물체추적

본 논문에서는 평행 이동 움직임만 검출하여 수행속도를 향상시킬 수 있는 투영모션 기법[5]을 사용한다. 그림 4는 2차원 영상에 대하여 가로, 세로로 각각 투영된 값들을 보여준다. 투영된 모션 기법은 각 방향에 따라 화

그림 4. 1차원 신호로 투영된 영상

소의 밝기 값을 누적하는 것으로 가로(x), 세로(y) 방향으로 1차원 신호의 투영정보를 만들어 비교영상에서 가로, 세로로 가장 유사한 영역을 찾는다. 그림 5는 비교영상에서의 탐색영역을 보이며 SR(Search Range)은 검색영역을 나타낸다.

2.3 템플릿 마스크

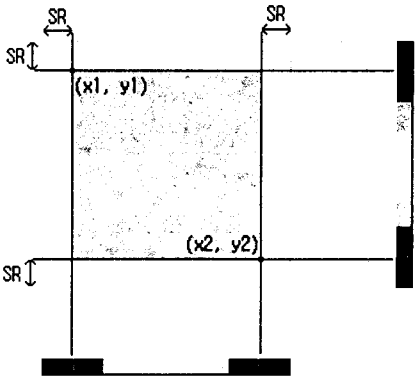

그림 5. 다음 프레임에서의 탐색영역

영역분할 및 객체추출 후 추출된 객체로 템플릿 마스크를 구성한다. 얻어진 템플릿 마스크를 스케일과 회전의 변화에 적용하기 위해 템플릿 마스크를 변화시킨다. 본 논문에서는 템플릿 마스크 회전 각도와 스케일 변화는 $\pm 10^{\circ}$ 와 $\pm 10\%$를 적용한다. 업데이트 된 템플릿 마스크는 다음 프레임에서 그대로 적용된다. 그림 6, 그림 7은 얻어진 템플릿 마스크와 회전된 템플릿 마스크를 보여준다.

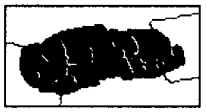

그림 6. 얻어진 템플릿 마스크

그림 7. 10도 회전된 영상

2.4 히스토그램 인터섹션

템플릿 마스크로 정의 된 질의영상과 프레임영상 간의 비교를 위해 채널당 3비트로 양자화 한 RGB 컬러 모델을 가지고 전체면적으로 정규화 한 확률 값으로 표현된 히스토그램 인터섹션 기법을 사용한다[7]. 식 (1)은 전체면적으로 정규화 한 확률 값으로 표현된 히스토그램 인터섹션 수식이다.

$$H(I,M)=\sum_{k=1}^{n}\min(\frac{I_j}{S_I},\frac{M_j}{S_M}) \quad (1)$$

여기서, I_j는 비교영상의 히스토그램 값, M_j는 질의영상의 히스토그램 값을 나타내며 S_I와 S_M 은 영상의 면적이다.

3. 실험 및 결과

실험을 위해 MPEG-1, MPEG-2의 동영상을 사용하였다. 표 1은 실험에 사용된 동영상을 보여준며 그림 8은 물체인식 결과영상을 보여준다.

표 1. 실험에 사용된 동영상 종류

동영상 종류	개수
광고	50
고속도로	20
도로	10
기타	20

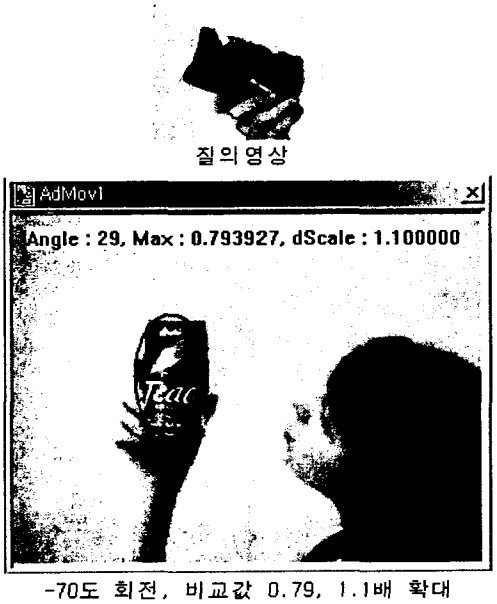

그림 8. 물체인식 결과영상

제안한 알고리즘을 비교하기 위해 Color Histogram Matching 기법과 Active Contour Model 기법[10]을 동일한 동영상에 적용하여 실험하였다. 그림 9는 실험에 사용된 동영상을 보여준다. 동영상 (A)는 객체이동, 동영상 (B)는 객체이동 및 스케일 변화, 동영상 (C)는 객체이동, 회전 및 카메라 모션이 발생하는 영상이다.

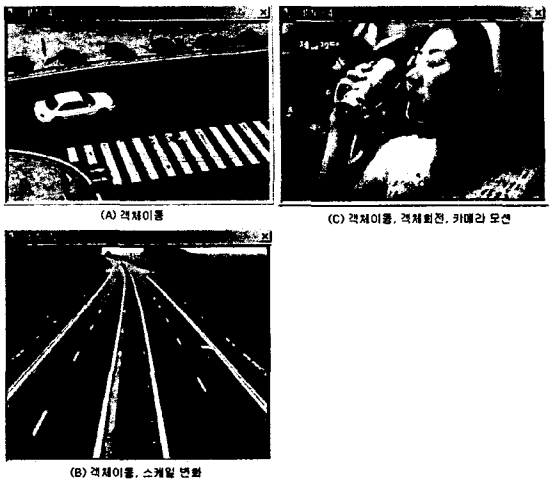

그림 9. 실험에 사용된 동영상

주위 검색영역(Search Area)은 모두 ±30 픽셀로 정의하였고 60프레임씩 비교하였다. 표 2, 그림 10은 실험동영상 (A), (B), (C)에 대한 수행시간을 비교한 결과이다.

표 2. 물체추적 수행시간 비교
(sec/frame)

동영상	제안된 방법	CHM 기법	ACM 기법
(A)	1.069	11.081	2.126
(B)	0.205	2.984	2.186
(C)	1.156	15.483	3.409

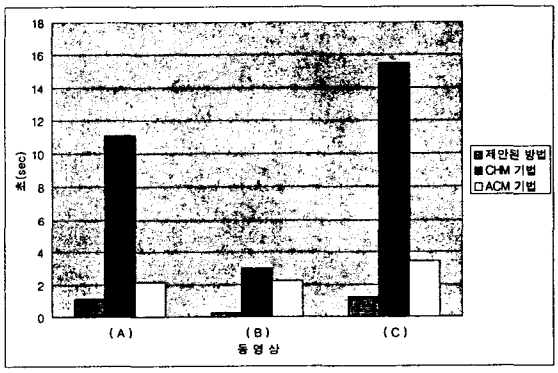

그림 10. 물체추적 수행시간 비교

표 3은 실험동영상 (A), (B), (C)에 대해 물체추적 성공여부 결과를 보여주며 표 4는 각 기법에 실험동영상을 적용한 후 테스트 한 결

과를 가지고 물체추적 기법에 대한 성능분석이다.

표 3. 물체추적 성공여부

동영상	제안된 방법	CHM 방법	ACM 방법
(A)	성공	성공	성공
(B)	성공	실패	성공
(C)	성공	실패	실패

표 4. 물체추적 기법들 성능 분석

비교 대상	제안된 방법	CHM 방법	ACM 방법
확대 및 축소	O	X	O
회전	O	X	O
비용 (시간)	보통	매우 많음	많음
객체 추적	좋음	보통	매우 좋음
카메라 모션	적용 가능	적용 가능	적용 불가

4. 결론 및 향후과제

본 논문에서는 투영된 모션과 히스토그램 인터섹션 기법을 이용한 노이즈에 강건한 물체추적 알고리즘을 제안하였다. 질의영상은 영상분할 후 영역선택을 통하여 구성하고, 물체추적은 가로 및 세로의 밝기 값을 1차원 신호로 투영하여 대략적인 움직임을 감지하며, 물체인식은 질의영상의 템플릿 마스크를 이동하여 회전과 스케일에 맞게 변경 한 후 감지하였다.

Color Histogram Matching 기법, Active Contour Model 기법과 비교해 본 결과 검색시간에서 우수한 결과를 보였고 객체의 회전 및 스케일 변화, 카메라 모션에도 적응적 임을 확인하였다.

향후 연구과제로는 배경과 객체를 자동으로 분리하는 방법과 템플릿 마스크를 사용하지 않고 질의영상의 특징을 보다 빠르고 효과적으로 나타낼 수 있는 방법의 연구가 가장 큰 과제이다. 동일한 배경에 동일한 객체가 여러 개 존재하는 경우, 객체의 한 부분이 다른 객체에 의해 가려졌을 경우, 객체가 화면 옆으로 사라지는 경우에도 추적이 가능하도록 연구되어져야 한다.

참고문헌

[1] "MPEG-4 Video Verification Model Version 13.0," ISO/IEC JTC1/SC29/WG11 MPEG99/N2687, 1999

[2] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 1992

[3] Randy Crane, "A Simplified Approach to Image Processing," Prentice Hall, pp. 206-211, 1997

[4] Jae Soo Seo, Jae Yeal Nam, Jin Suk Kwak, Myoung Ho Lee, "Adaptive Predicted Direction Search Algorithm for Block Matching Motion Estimation," 제12회 영상처리 및 이해에 관한 워크샵 발표 논문집, pp.415-420, 2000

[5] Dong Kwon Park, Ho Seok Yoon, Woo Sung Jeon, Chee Sun Won, "Semi-Automatic Object Tracking Algorithm Using Projected Motion," 제12회 영상처리 및 이해에 관한 워크샵 발표 논문집, pp.139-144, 2000

[6] Jung Bum Oh, Young Shik Moon, "Content-Based Image Retrieval Based on Scale-Space Theory," IEICE Trans. Fundamental, June, 1999

[7] Anil K. Jain and Aditya Vailaya, "Image Retrieval Using Color and Shape," Pattern Recognition, vol.29, no.8, pp.1233-1244, 1996

[8] Chad Carson, Serge Belongie, Hayit Greenspan, and Jitendra Malik, "Region -Based Image Querying," CVPR '97 Workshop on Content-Based Access of Image and Video Libraries, 1997

[9] Demin Wang, "Unsupervised Video

Segmentation Based on Watersheds and Temporal Tracking," *IEEE Trans. Circuits and System for Video Technology.*, vol. 8, no. 5, pp.539-546, Sep., 1998

[10] Nikos Paragios, Rachid Deriche, "Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Objects," *IEEE Trans. Pattern Analysis and Machine Intelligence.*, vol. 22, no. 3, pp.266-280, Mar., 2000