

실시간 영상에서 객체 추적 및 얼굴추출

이광형*, 김용균*, 지정규**, 오해석*

*승실대학교 컴퓨터공학과

**한국학술진흥재단

e-mail:loke7777@yahoo.co.kr

Object Tracking and Face extract by Real-time Image

Kwang-Hyoung Lee* Yong-Gyun Kim* Jeong-Gyu Jee** Hae-Seok Oh*

*Dept. of Computing, Graduate School, Soongsil University

**Korea Research Foundation

요 약

실시간 영상에서 객체 추적은 수년간 컴퓨터 비전 및 여러 실용적 응용 분야에서 관심을 가지는 주제 중 하나이다. 실제로 실시간 영상내의 객체 추적은 빠른 처리와 많은 연산은 요구하고 고가의 장비가 필요하기 때문에 많은 어려움이 따른다. 본 논문에서는 보안시스템에 적용될 수 있게 실시간으로 배경영상을 갱신하면서 객체를 추출 및 추적하고 추출된 객체에서 얼굴을 추출하는 방법을 제안한다. 배경영상과 입력영상의 차이를 이용하여 실시간으로 배경영상을 입력영상으로 대체하여 시간의 흐름에 의한 배경잡음을 최소화 하도록 적응적 배경영상을 생성한다. 그리고 배경영상과 카메라로부터 입력되는 입력영상과의 차를 이용하여 객체의 크기와 위치를 탐지하여 객체를 추출한다. 추출된 객체의 내부점을 이용하여 최소사각영역을 설정하고 이를 통해 실시간 객체추적을 하였다. 또한 설정된 최소사각영역은 피부색의 RGB 영역에서 얼굴 영역을 추출하는 데도 적용한다.

1. 서론

실시간 영상에서 객체 추적을 위해 컴퓨터 비전 기술을 이용하여 구현하기란 매우 어려운 일임에도 불구하고, 컴퓨터 성능의 발달로 인해 영상 처리 기법의 발전과 더불어 객체 인식과 객체 추적에 관한 연구가 활발히 진행되고 있다. 객체를 추적하는 방법은 보안, 의료, 군사, 교통, 제어분야 등 여러 분야에 응용될 수 있어 그동안 많은 연구와 개발이 이루어지고 있다[1].

컴퓨터 비전 시스템에서 사용되는 객체 추적 방법은 각각 특정 환경에서 적합한 동작을 보이도록 몇 가지 제약 조건을 두어 설계되고 있다[2]. 본 논문에서는 보안 / 감시 시스템 분야에서 적용되어 질 수 있는 방법을 제안하기 위하여, 고정형 저가형 PC카메라를 사용하였고, 실내환경의 제약 조건을 가진다.

본 논문은 시간의 경과에 따라 배경영역에서 잡음 생성을 줄이고자 객체영역 이외의 영역을 배경영역

으로 갱신하여 항상 최신의 배경영상을 유지하도록 하였다. 객체의 추출은 생성된 적응적 배경영상과 실시간으로 입력되는 입력영상의 차를 이용하여 객체의 크기와 위치를 탐지하여 그물식 탐색방법으로 객체를 추출한다. 추출된 객체의 내부점들을 MBR(Minimum Bounding Rectangle)을 설정하여 객체의 실시간 추적을 가능하도록 하였고 MBR내에서 RGB 영역의 색상 정보를 이용하여 얼굴영역을 추출하기 위한 방법으로 사용한다. 본 논문은 실시간으로 입력되는 영상으로부터 적응적 배경영상으로 잡음을 현저히 줄였고 실시간으로 얼굴 영역을 추출함으로써 보안 / 감시 시스템으로 효용성을 향상시켰다.

2. 관련연구

기존의 추적 알고리즘은 크게 특징기반방법(feature-based approaches)과 이미지기반방법(image-based approaches)으로 나누어진다. 특징기

반방법은 저수준 해석(Low-Level Analysis), 특징 해석(Feature Analysis), 능동 형상 모델(Active Shape Models) 방법이 있고, 이미지기반방법에는 선형 부분 공간 방법(Linear Subspace Methods), 신경망(Neural Networks), 통계적 방법(Statistical Approaches) 등이 있다.

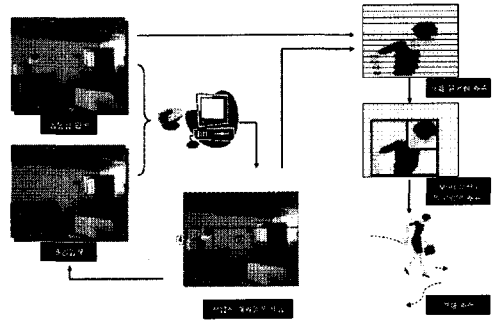
특징 해석은 저수준 해석으로부터 생성되는 특징들은 애매하기 십상이다. 예를 들어서 컬러 모델을 사용하여 객체 영역들의 위치를 결정하는 것에서 유사한 컬러의 배경 객체들 또한 검출될 수 있다. 이것은 더 높은 수준의 특징해석에 의해서 해결될 수 있다. 많은 객체검출 기법들에서 객체형태의 지식은 그것들의 애매한 상태로부터 다양한 특징들을 특성화하고 이어서 입증하기 위해서 채택되었다. 조사된 문헌들에서는 객체형태의 응용과 관련한 두 가지 방법들이 존재한다. 첫 번째 방법은 개인의 안면 특징들의 상대적인 위치결정에 기초하는 순차적 특징 탐색 전략들을 포함한다. 두 번째 방법그룹에 있는 기법들은 다양한 얼굴모델들을 사용하는 유연한 배열들을 특징으로 한다[4]. 능동 형상 모델들은 특징들의 실제적인 물리적 및 더 높은 수준의 외관을 묘사한다. 일단 특징에 대해 가까운 근접 내에서 방출되면 능동 형상 모델은 국소 이미지 특징들(모서리들, 밝기)과 상호작용하고 점차적으로 특징의 형상을 얻기 위해서 변형할 것이다. 현대의 객체 추출연구에는 일반적으로 세 가지 종류의 능동 형상 모델들이 존재한다.[5]. 사람얼굴은 효과적으로 표현하기 위해 개발된 PCA방법은 다른 얼굴 이미지들의 조합이 주어지면 고유벡터들에 의하여 표현되는, 얼굴들의 분포의 주성분들을 찾는다. 그 후에 얼굴집합에 있는 각 개별 얼굴이 적절한 가중들을 사용하여 더 일반적으로는 고유 얼굴들로 나타내지는, 가장 큰 고유벡터들의 선형조합에 의해서 근사될 수 있다.[6].

3. 실시간 객체 추적 및 얼굴영역 추출

3.1 시스템 구조

컴퓨터 비전 시스템을 이용하여 객체를 추적하는 방법은 각각 특정 환경에서 적합한 동작을 보이도록 몇 가지 제약 조건을 두어 설계되어져 왔다. 센서 장착이나 움직임이 가능한 고가의 카메라가 아닌 저가형 PC카메라를 이용하는데 이는 카메라의 움직임이 고정됨을 의미한다. 또한 사무실이나 연구실, 기자재실, 자재창고, 주차장, 은행 365일 코너 등 배경영상의 변화가 미미한 특수 환경으로 환경을 제한한

다. [그림 1]은 제안하는 방법의 전체적인 처리 흐름도이다. PC카메라로 실시간 받아들여지는 영상에서 객체의 출현이 감지되지 않는 영상을 초기의 배경영상으로 설정하고 연속적으로 입력되어지는 영상과 초기 배경영상의 비교를 통하여 강인한 객체 추출을 할 수 있는 적응적 배경영상을 생성한다.



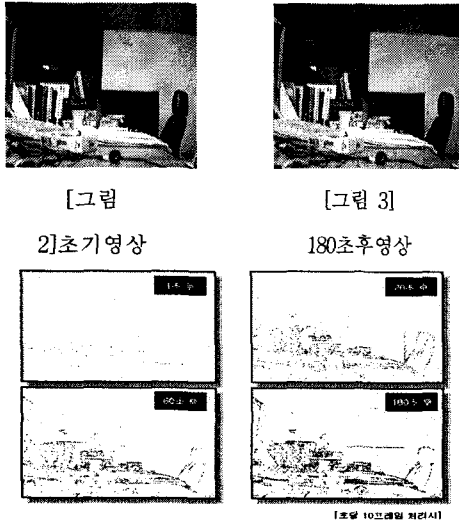
[그림 1] 시스템 처리 흐름도

계속적으로 갱신을 반복하는 배경영상과 실시간 입력되는 영상에서 그물식 탐색 방법을 이용하여 객체의 출현을 감지하고, 감지된 객체의 내부점을 추출한다. 추출된 객체 내부점을 이용하여 객체를 포함하는 최소 사각형인 MBR을 설정해 주고, 연속되어지는 영상으로부터 일련의 처리를 반복하게 된다. 연속된 MBR의 설정을 추적함으로써 객체의 실시간 추적을 가능토록 하였다. 추적과 동시에 전체 영상이 아닌 MBR 내에서 얼굴 영역을 추출하는 방법을 사용하여 실시간 얼굴영역을 추출한다.

3.2 시스템 구현 방법

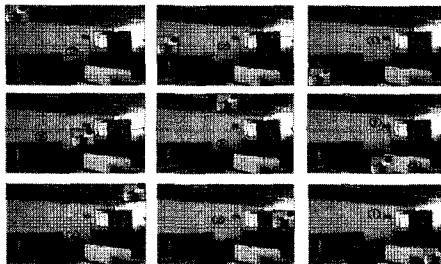
3.2.1 적응적 배경영상 생성

[그림 2]와 [그림 3]은 PC카메라로부터 입력되어진 초기 설정된 배경영상과 객체의 움직임이 감지되지 않은 상태에서 180초 후의 배경영상이다. 사람의 감지 능력으로는 두 그림이 똑같아 보이지만, 시스템에서는 둘을 다르게 인식한다. [그림 4]가 시스템이 두 영상을 다르게 인식함을 보이는데, 어떠한 변화도 없는 배경영상이 조명의 영향과 시간의 경과에 따른 잡음 발생의 현상을 보인다. 1초, 20초, 60초, 180초 후 영상에서 각각 71개, 1657개, 2061개, 3891개의 최초 배경영상과의 차에서 임계값을 넘어가는 픽셀의 개수가 나타난다. 전체적인 배경영상의 갱신은 많은 연산량을 필요로 하기 때문에 실시간 객체 추적 성능을 저하시킨다. 제안하는 시스템에서는 객체영역인 $n * m$ 의 영역을 제외한 나머지 영역을 배경영상으로 갱신 하면서 객체 추적을 동시에 이루어질 수 있도록 하였다.



[그림 4] 시변하는 배경영상에서의 잡음

[그림 5]는 적응적 배경영상 생성 방법을 보인다. 이전의 배경영상과 입력영상에서 객체 추출 후 객체 영역을 제외한 나머지 영역을 새로운 배경영상으로 대체한다. [그림 5]는 객체의 위치에 따른 취할 수 있는 배경영상의 집합이다. 객체의 위치에 따라 ①~④ 부분으로 분류하고 입력영상을 배경영상으로 대체한다.



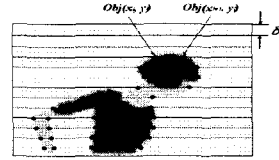
[그림 5] 적응적 배경영상 생성

제안한 적응적 배경영상 방법을 이용하여 실험한 결과, 시변하는 배경영상내의 조명 등에 의한 잡음을 객체 내부에 포함되어 있는 배경영역을 제외한 나머지 영역에서 95% 이상 제거할 수 있었다.

3.2.2 객체의 내부점 추출

객체의 추적을 위해서 객체의 위치 결정을 위한 객체 추출이 우선적이다. 제안하는 시스템에서는 그물식 탐색 방식을 이용한다. 그물식 탐색 방식은 영상을 위에서 아래로 δ 픽셀 간격으로 탐색해 나가는 방법으로 완전 차영상을 이용하는 시스템보다 처리 속도의 높은 향상을 보인다. [그림 6]은 그물식 탐색 방식 및 객체 내부점 추출이다. (식 1)에 따라 RGB

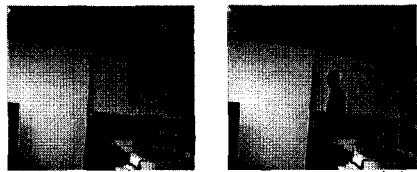
채널 각각의 차이가 임계값 β 보다 크다면, 객체 내부점 후보군으로 등록하게 된다. (식 2)는 잡음제거를 위한 식이다. (식 1)를 만족하는 객체 내부점 후보군에서 연속된 픽셀이 γ 이하라면 잡음으로 판단하여 제거하고, 원하는 객체 내부점 $Obj(x_i, y_i)$ 를 추출하게 된다.



[그림 6] 그물식 탐색 방식

$$\begin{aligned}
 R_{obj} &= \text{abs}(RV(BG[x_i, y_i]) - RV(TG[x_i, y_i])) \\
 G_{obj} &= \text{abs}(GV(BG[x_i, y_i]) - (GV(TG[x_i, y_i]))) \\
 B_{obj} &= \text{abs}(BV(BG[x_i, y_i]) - (BV(TG[x_i, y_i]))) \\
 \text{if } ((R_{obj} \geq \beta) \text{ or } (G_{obj} \geq \beta) \text{ or } (B_{obj} \geq \beta)) \\
 &\quad ObjInit(x_i, y_i) = \text{true} \\
 \text{else } &\quad ObjInit(x_i, y_i) = \text{false} \quad (\text{식 1}) \\
 &\quad * ObjInit(x_i, y_i) : \text{객체 내부점 후보군} \\
 &\quad * \beta : \text{객체 내부점 후보군 추출 임계값} \\
 x_{i+1} - x_i &\geq \gamma \quad (\text{식 2}) \\
 &\quad * \gamma : \text{객체 내부점 추출 임계값}
 \end{aligned}$$

[그림 7]은 계속적으로 갱신을 반복하는 적응적 배경영상과 실시간 입력되어지는 영상을 나타낸다. 배경영상과 입력영상에서 그물식 탐색 방식을 이용한 객체 내부점 추출을 [그림 8]이 나타낸다.



[그림 7] (a) 배경영상, (b) 입력영상



[그림 8] 객체 내부점 추출

3.2.3 MBR을 이용한 객체 추적

추출된 객체 내부점을 이용하여 MBR(Minimum Bounding Rectangle)을 설정하여 준다. MBR은 객체 추적을 위한 객체를 포함하는 최소 사각 영역의

로써, 얼굴 영역을 추출할 때 속도를 향상시키기 위한 범위를 제한해 주는 역할을 수행한다. 객체 내부점 $Obj(x_i, y_i)$ 의 x, y 좌표 중 (식 4)를 통하여 각각의 최대, 최소 좌표를 구하여 객체를 포함하는 최소한의 사각 영역을 설정한다. [그림 9]는 추출된 객체 내부점을 이용하여 설정된 MBR을 보인다. 이를 통하여 객체의 실시간 추적을 가능하도록 하였다.

$$Obj_left = \min[Obj(x_i)] - \delta, Obj_top = \min[Obj(y_i)] - \delta$$

$$Obj_right = \max[Obj(x_i)] + \delta, Obj_bot = \max[Obj(y_i)] + \delta$$

$$MBR = [Obj_left, Obj_top, Obj_right, Obj_bot] \quad (식 4)$$

* $Obj(x_i, y_i)$: 객체내부점



[그림 9] 설정된 MBR

3.2.4 얼굴 영역 추출

정규화된 RGB를 이용한 방법은 정규화된 값을 사용하여 HSI와 유사한 효과를 거뒀다. 이때 $r'+g'+b'=1$ 이라는 성질을 이용하여 r', b' 의 두 채널 정보만을 사용하여 색을 구분하는 기준으로 사용할 수 있다. RGB 좌표계에서 색상의 정규화 식은 (식 5)과 같다. 이 방법은 R, B의 값만을 이용하여 얼굴 영역 검출에 이용할 수 있다.

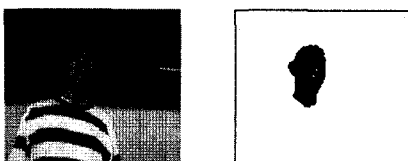
$$(r', g', b') = \left(\frac{r}{r+g+b}, \frac{g}{r+g+b}, \frac{b}{r+g+b} \right) \quad (식 5)$$

정규화된 RGB 모델 실험에서 피부색의 범위를 다음과 같이 사용하였다.

$$r'_{\min} = 0.3550, r'_{\max} = 0.4300$$

$$b'_{\min} = 0.2500, b'_{\max} = 0.3050$$

이 방식은 어두운 영역에서 후보 검출에 장점이 있으나 유사 색상영역에서는 구별이 어렵다. 이에 본 논문에서는 두 가지 방법을 혼합한 Hybrid 방법을 이용하였다. [그림 10]는 입력되어진 영상에서 색상계를 이용한 얼굴영역검출 결과를 보이고 있다.



[그림 10] 얼굴 영역 검출

4. 결론 및 향후 연구방향

본 논문에서는 입력되어지는 영상으로부터 배경영상의 실시간 갱신을 통해 객체의 위치를 탐지하고 객체를 추적하고, 객체의 내부점을 이용한 MBR에서의 얼굴영역을 추출하는 방법을 제안하였다. 실험은 고정 PC카메라와 배경영상의 변화가 거의 없다는 제한된 환경 조건에서 실시간으로 배경영상의 갱신과 객체의 추적이 안정적인 모습을 보여주었다. 이는 객체의 추출 및 추적 알고리즘이 빠르게 수행되어 객체 인식과 결합하여 객체의 움직임 정보와 인식을 통한 보안 / 감시 시스템 등 응용분야에 적용되어질 수 있다는 것을 기대할 수 있도록 하여 준다. 하지만 입력영상과 배경영상 각각의 R, G, B 값의 차를 이용, 객체의 일부분임을 인식할 때, 잡음과 조명의 영향으로 인하여 실험에 실패한 경우도 발생하여 이에 대한 보완이 필요하다. 그리고 실시간으로 입력되어진 영상에서 얼굴영역을 추출하여 추출된 얼굴의 정면에 가장 가까운 영상을 추출해 내는 연구가 진행 중이다.

참고문헌

- [1] 이희영, 최재영, 강동구, 김홍수, 차의영, 전태수, "배경영상을 이용한 목표물 추적에 관한 연구", 한국멀티미디어학회 1999년도 춘계학술발표논문집 (학술발표), Vol.2, No.1, pp.386-390, 1999.
- [2] 황본우, 손형진, 이성환, "대화형 하이퍼 비디오 저작을 위한 객체 기반 추적 방법", 정보과학회 2001년 추계학술대회, Vol.28, No.2, pp.427-429, 2001.
- [3] E. Hjelm and B.K. Low, "Face Detection: A Survey." Computer Vision and Image Understanding, Vol.83, No.3, pp.236-274, 2001.
- [4] T.V. Pham, M. Worring, and A.W.M. Smeulders, "Face detection by aggregated Bayesian network classifiers." Technical Report 2001-04, Intelligent Sensory Information Systems Group, University of Amsterdam, 2001.
- [5] M. Rogers and J. Graham, "Robust active shape model search." Proceedings of the 7th European Conference on Computer Vision, No.4, pp.517-530, 2002.
- [6] K. Okada and C. von der Malsburg, "Analysis and synthesis of human faces with pose variations by a parametric piecewise linear subspace method." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, No.1, pp.761-768, 2001.