

칼라 히스토그램을 이용한 비디오 영상에서 얼굴 영역의 고속 추적

유 태 응, 오 일 석
전북대학교 컴퓨터과학과

Fast Tracking of Face Region in Video Images using Color Histogram

Tae-Woong Yoo and Il-Seok Oh
Dept. of Computer Science, Chonbuk National University

요 약

본 논문은 비디오 연속 영상에서 얼굴의 위치를 추적하는 알고리즘에 관하여 기술한다. 컴퓨터 비전에서 대량의 비디오 연속 영상내 물체 추적은 실시간에 처리되는 빠른 알고리즘이 요구된다. 기존의 방법은 형태에 기반한 알고리즘으로 물체의 회전, 크기 변화, 겹침 등에 대한 문제에 민감하여 여러 가지 어려움이 발생한다. 그러나 칼라를 이용한 알고리즘은 이러한 문제에 대하여 둔감하여 훨씬 효과적이다. 본 논문은 칼라 3D 히스토그램을 이용한 Swain과 Ballard의 역 투사(backprojection) 방법을 적용하여 비디오 연속 영상에서 얼굴의 위치를 빠르고 정확히 추적하는 알고리즘을 제안한다.

1. 서 론

컴퓨터 비전에서 물체를 추적하는 기존의 알고리즘 대부분은 명암 영상을 가지고하는 광류 정보, 형태(template matching, edge detection 등)정보에 기반한 알고리즘들이 주를 이룬다[1]. 물체 추적 알고리즘으로 널리 사용되는 광류 기반 알고리즘은 각 프레임 내의 각 화소의 변위벡터나 속도벡터를 이용하여 물체를 추적하는 알고리즘이다. 광류 기반 알고리즘은 많은 계산량과 잡음에 민감하여 비디오 연속 영상에서 실시간 얼굴 추적에는 부적합하다. 형태 기반 알고리즘[2, 3]은 물체의 회전, 물체의 크기 변화, 물체의 겹침, 그리고 잡음에 대하여 매우 민감하여 연속 영상에서 얼굴 추적시에 많은 어려움이 나타난다. 칼라 정보의 사용은 이러한 문제점들을 둔감시킬 수 있다. 특히, 비디오 연속 영상에서 얼굴 위치를 추적할 때, 얼굴 칼라 분포는 대부분 유일하고 프레임 영상 내의 얼굴의 모양 변화 및 이동, 크기, 회전 등의 변화에 둔감하여 쉽게 얼굴 위치를 추적할 수 있다. 그리고 프레임 영상 내의 특정 칼라의 검색은 단순한 계산에 의해 수행될 수 있어서 실시간 처리가 가능하다.

본 논문은 칼라 정보(3D 히스토그램)를 이용한 고속

얼굴 추적 알고리즘을 제안한다. 칼라 3D 히스토그램은 현재 프레임 영상의 각 화소의 R, G, B 요소 값이 산화하여 해당 bin을 증가시켜 계산한다. 구해진 히스토그램은 다음 프레임에 Swain-Ballard의 알고리즘[4]을 사용하여 역 투사된다. 그리고 역 투사된 영상에서 이전 얼굴 위치를 중심으로 마스크를 적용하여 빠르게 현재 프레임에서 얼굴 위치를 찾는다. 이때 탐색은 나선형으로 수행하여 탐색 공간을 제한하여 계산 효율을 높인다. 이러한 칼라의 사용은 물체 추적의 빠른 수행 능력, 관점 변화에 대한 처리 능력, 물체의 변형과 배경으로부터 부정확한 segmentation 등을 처리할 수가 있다. 실험 결과 비디오 연속 영상(160×120 / frame)에서 초당 2-3 프레임은 정확하게 처리한다.

본 논문의 구성은 다음과 같다. 2 절은 얼굴 추적 알고리즘에 대하여 기술하고 3 절과 4 절은 실험 결과의 향후 연구 과제에 대하여 기술한다.

2. 얼굴 추적 알고리즘

2.1 히스토그램 배경 투사

Swain과 Ballard의 역 투사 알고리즘은 3D 칼라 히

스토그램을 기반으로한 알고리즘으로 영상 전체의 통계적인 정보에 의존하며 point-to-point 처리를 한다[4]. 모델 히스토그램(M)과 영상 히스토그램(I)을 가지고 비율 히스토그램(R)을 구한다. 즉, 추출된 칼라의 특성이 얼마 만큼인가를 표현하는 값을 가지고 프레임 영상 내의 각 칼라에 대치시킨다. 영상 내의 칼라 i 의 화소 (x, y) 는 M_i / I_i 값으로 대치된다. M_i 는 모델 히스토그램의 bin i 내의 갯수, I_i 는 영상 히스토그램의 bin i 내의 갯수이다. Swain과 Ballard의 히스토그램 칼라 공간과 역 투사 알고리즘은 다음과 같다.

- 히스토그램 칼라 축은 다음과 같은 3개의 칼라 축을 사용한다[1].

$$\begin{aligned}rg &= r - g \\by &= 2 * b - r - g \\wb &= r + g + b\end{aligned}$$

- 역 투사 알고리즘

① 각 히스토그램 bin i 에 대하여

$$R_i = \min \left[\frac{M_i}{I_i}, 1 \right]$$

② 함수 $h(c)$ 는 한 칼라 c 를 히스토그램에 사상시킨다. 각 x, y 에 대하여

$$b_{x,y} = R_{h(c_{x,y})}$$

③ $b = D' * b$,

$$D'_{x,y} = \begin{cases} 1 & \text{if } \sqrt{x^2 + y^2} < r \\ 0 & \text{otherwise} \end{cases}$$

④ loc는 역 투사된 영상 전체 내에 원형 마스크를 적용하여 원형 마스크 내부의 값들을 더하여 가장 큰값을 갖는 위치를 말한다. 각 x, y 에 대하여

$$(x_l, y_l) = \text{loc}(\max_{x,y} b_{x,y})$$

2.2 추적 알고리즘

얼굴 영역은 피부색 분포를 이용하여 첫 프레임에서 추출한다[5, 6]. 추출된 모델 영역의 각 화소의 R, G, B 요소 값을 가지고 3D 칼라 히스토그램(모델 히스토그램)을 구한다. 현재 프레임 영상에서 각 화소의 R, G, B 요소가 속하는 bin을 증가시킨 후 Swain과 Ballard 역 투사 알고리즘을 이용한다. 그러나 Swain과 Ballard 방법에서는 프레임 영상에 역 투사되는 값이 1 아니면 실수가 투사되어 계산 값의 범위가 넓어지고 투사된 영상에서 원형의 마스크를 프레임 영상 전체에 적용하여 계산하는데 많은 시간이 소비된다. 그러므로 계산량을 줄이기 위하여 역 투사되는 값을 0 아니면 1을 주어 이전 영상을 획득한다.

획득된 이전 영상을 얼굴 위치 검색을 빠르게 하기 위하여 피라미드 기법(pyramid technique)[1]을 사용하여 피라미드 영상으로 사상(mapping)한다. 사상된 피라미드 영상에서 얼굴의 모양과 비슷한 타원 마스크를 이용하여 얼굴 위치를 찾는다. 추출된 얼굴 위치 좌표를 원 프레임 영상으로 사상시킨다. 이 얼굴 위치 좌표를 가지고 프레임 연속 영상에서 얼굴 칼라가 환경에 의해 변하여도 그 다음 프레임 영상에서 얼굴 위치를 찾기 위하여 다시 새로운 모델 히스토그램을 구한다. 그리고, 비디오 연속 영상에서 얼굴의 위치는 갑자기 크게 이동하지 않으므로 그 다음 프레임에서 이전 프레임의 얼굴 위치 중앙 좌표를 중심으로 나선형 탐색을 한다. 나선형 탐색을 함으로써 연속된 프레임에서 얼굴 위치를 추출하는데 정확성을 보장한다. 제안한 얼굴 추적 알고리즘은 다음과 같다.

① 각 히스토그램 bin i 에 대하여

$$R_i = \begin{cases} 1 & \text{if } \frac{M_i}{I_i} \geq 1 \\ 0 & \text{if } \frac{M_i}{I_i} < 0 \end{cases}$$

② 각 x, y 에 대하여 $b_{x,y} = R_{h(c_{x,y})}$

③ 각 x, y 에 대하여 $b_{x,y} \rightarrow p_{x,y}$ 로 사상
for($y=0; y < y_{coor}, y+=2$)

for($x=0; x < x_{coor}, x+=2$) {

if ($(b_{x,y} + b_{x,y+1} + b_{x+1,y} + b_{x+1,y+1}) \geq 2$)

$$p_{(y_{coor}/4 + y/2), (x_{coor}/4 + x/2)} = 1$$

}

④ p 에서 이전에 찾아진 얼굴 위치를 중심으로 나선형 검색

$$p = D' * p,$$

$$D'_{x,y} = \begin{cases} 1 & \text{if } \frac{x^2}{a^2} + \frac{y^2}{b^2} < 1 \\ 0 & \text{otherwise} \end{cases}$$

⑤ $(x_p, y_p) = \text{loc}(\max_{x,y} p_{x,y})$

⑥ 피라미드 영상의 얼굴 위치 좌표를 원 프레임 영상 위치로 사상 : $p_{x,y} \rightarrow (x_l, y_l)$

$$(x_l, y_l) = \begin{cases} x_l = (x_p - x_{coor}/4) * 2 \\ y_l = (y_p - y_{coor}/4) * 2 \end{cases}$$

⑦ 새로운 모델 히스토그램 M_i 를 구한다.

3. 실험 결과

본 실험은 X 라이브러리를 사용하여 TG-SDT500 워크스테이션에서 구현하였다. 비디오 연속 영상의 한 프레임 크기는 160×120으로 약 1000 프레임 정도의 연속 영상을 초당 2-3 프레임 처리하였다. 다음 그림 1은 실험 과정 중 중간 단계를 보여주는 비디오 연속 영상의 일부이다.

루어야 한다. 또한 얼굴 영역내에서 입, 눈 같은 특징을 찾아 추적하는 문제도 해결되어야 한다.



그림 1 실험 과정을 보이는 비디오 연속 영상

그림 1의 첫 번째 행은 원 프레임 연속 영상이다. 두 번째 행의 연속 영상은 원 연속 영상에서 이진 영상을 구한 것이다. 이진 영상에서 하얀 칼라의 밀도가 높을수록 얼굴 위치를 암시한다. 그 다음 세 번째 행의 연속 영상은 타원 마스크를 적용하였을 때 계산 속도를 높이기 위하여 이진 영상을 피라미드 영상으로 사상한 연속 영상이다. 피라미드 연속 영상에서 타원 마스크를 이용하여 얼굴 위치를 추출한다. 마지막 행의 연속 영상은 피라미드 연속 영상에서 추출된 얼굴 위치를 원 프레임 연속 영상에 사상시켜 얼굴 위치를 표시한 프레임 연속 영상이다. 다음 그림 2는 얼굴 추적 연속 영상을 보인다.

4. 향후 연구 과제

본 연구에서는 비디오 연속 영상에서 얼굴 추적 알고리즘을 제안하였다. 실험을 통하여 초당 2-3 프레임을 처리 하였다. 이는 실시간 처리에 부족한 속도이지만 빠른 컴퓨터를 사용하면 쉽게 실시간 처리가 가능할 것이다. 여러 사람의 얼굴을 동시에 추적하는 문제로 다

참 고 문 헌

- [1] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, Addison-Wesley, 1993.
- [2] L. C. Desilva, K. Aizawa and M. Hatori, "Detection and tracking of facial features," *SPIE Visual Communications and Image Processing '95*, Vol. 2501, 1995, pp. 1161-1172.
- [3] S. Katahara and M. Aoki, "Motion tracking of eyes and mouth by extracting the contours," *Asian Conference on Computer Vision '93*, November, 1993, pp. 363-366.
- [4] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal Computer Vision*, Vol. 7, No. 1, pp. 11-32, 1991.
- [5] T. C. Chang and T. S. Huang, "Facial feature extraction from color Images," *Proceeding of 12th International Conference on Pattern Recognition*, Vol. 2, 1994, pp. 39-43.
- [6] T. W. Yoo and I. S. Oh, "A method for face

region tracking in real time based on color histogram," *SPIE Conf*(Submitted).



그림 2 얼굴 추적 실험 영상