

특징 추출과 변형가능 템플리트를 이용한 동영상에서의 얼굴 트래킹

Face Detection Using Feature extraction and Deformable Template in Motion Images

위 성윤, 윤 장용, 지 승환, 박 민용

연세대학교 전자공학과
 서울 서대문구 신촌동 134 120 - 749
 E-mail: wsy@yecis.yonsei.ac.kr

요 약

본 논문에서는 CCD 카메라로부터 획득한 영상 시퀀스들에서 인접한 두 영상사이의 차영상과 얼굴이 가지는 컬러정보를 이용하여 분리한 얼굴 영역에서 양쪽 눈과 입의 위치좌표를 특징점으로 이용하여 입력된 얼굴에 자동정합된 변형가능 템플리트(Deformable Template)를 가지고 연속된 다음 프레임에서 얼굴 전체를 트래킹하는 알고리즘을 제안한다. 실제 입력영상의 얼굴 영역과 변형가능 템플리트의 차이를 비교하기 위해 텍스처 매핑(Texture mapping)을 도입하여 트래킹의 정확도를 살펴본다.

1. 서 론

정지영상이나 동영상으로부터 다수의 얼굴을 인식하거나 한 개인의 얼굴을 식별하는 기술은 영상 처리나 패턴 인식, 컴퓨터 비전과 신경망 등과 같은 다양한 분야에서 활발히 연구되어지고 있으며, 상업적으로 법적으로 수 많은 응용 사례를 가지고 있다.

이런 얼굴 인식 기술(Face Recognition Technology)은 여권, 신용 카드, 주민 등록증, 범인의 얼굴 사진 등과 같은 제한된 형태의 정지 영상을 사용하는 방법에서부터 동영상 감시 시스템 등과 같은 실시간 인식에 이르기까지 그 종류가 다양하다.

얼굴 인식 문제는 3 가지로 나누어 질 수 있다. 첫째, 입력된 영상에서 얼굴을 분리하는 문제, 둘째, 특징점을 추출하는 문제, 마지막으로 얼굴을 인식하거나 분류하는 문제로 나누어진다. 본 논문에서는 위의 세 가지 과정 중에서 얼굴 영

역 분리와 분리한 얼굴 영역에서의 특징점 추출에 기반한 변형가능 템플리트를 이용한 얼굴 트래킹을 구현한다.

템플리트를 사용해 얼굴을 찾거나 얼굴의 특정 부분을 트래킹하는 방법의 경우에는 패러미터(parameter)들의 복잡한 계산과 많은 시간 지연이 발생하였다. 하지만, 본 논문에서는 입력 영상의 얼굴 영역에서 특징점인 양쪽 눈의 위치를 찾아내어 템플리트를 실제 얼굴의 형태와 방향을 가지도 록 구현한 프로그램을 사용하여 이러한 단점을 보완한다.

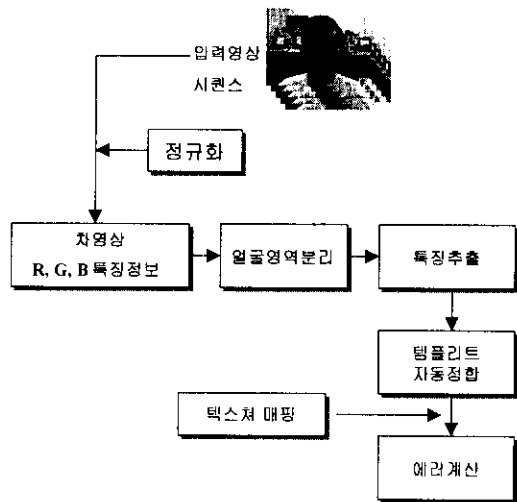


그림 1. 얼굴 트래킹 과정

그림 1은 본 논문에서 제안한 얼굴 트래킹 시스템의 구성도이다. 입력으로는 320 × 240 크기의 컬러 RGB 영상 시퀀스들을 획득한다. 이 획득된

입력 영상을 정규화 시킨 후, 두 인접한 영상사이의 차영상과 얼굴이 가지는 독특한 R, G, B 값을 사용하여 얼굴을 분리해 낸다. 분리된 얼굴영역에서 양쪽 눈의 위치정보를 구한 후, 이 위치정보를 바탕으로 양쪽 눈과 입의 삼각비를 이용해 표준템플리트를 입력영상의 얼굴에 맞도록 자동정합시킨다. 이 정합된 템플리트를 사용해 계속된 영상 시퀀스에서 얼굴을 트래킹한다. 이 변형가능 템플리트는 그리드(grid)간의 연결된 다수의 삼각형으로 구성되며, 각각의 그리드 내부에 x, y 좌표의 위치정보를 포함하게 된다. 그러나, 실제 입력 영상과 트래킹하는 변형가능 템플리트 사이의 직접적인 비교가 불가능하다. 따라서, 얼굴 트래킹의 정확도를 고찰하기 위해서 실제 영상에서 20 대 남자의 표준얼굴을 텍스처 매핑시킨 결과 영상의 차를 이용해 트래킹의 정확도를 파악한다.

이 식(3)을 통해 인접한 두 프레임사이의 차영상이 구해진다. 이 구해진 차영상에서 얼굴이 가지는 독특한 R, G, B 값들을 적용하여 실제 얼굴영역을 구하게 된다. 그림 2 는 시간 t0 와 t0 + T의 영상에서 차영상과 R, G, B 특성정보를 바탕으로 찾아낸 얼굴 영역을 나타내고 있다.

2. RGB 좌표계의 정규화

본 논문에서 획득한 입력 영상은 24bit RGB 컬러영상이다. 컬러 영상은 조명(illumination)에 매우 민감한 영향을 받게 되며 조명에 대한 영향을 고려하지 않는다면 실제 영상처리과정에서 많은 오류가 발생할 수 있다. 또한, 시간에 따른 조명의 변화는 비전 시스템의 성능에 많은 영향을 끼치게 됨을 경험적으로 알 수 있다. 본 논문에서는 이러한 문제점을 해결하기 위해 아래와 같은 정규화 과정을 통하여 획득한 RGB 컬러영상의 조명에 대한 의존도를 감소시킨다.

$$D = R + G + B \quad (1)$$

$$r = \frac{R}{D}, \quad g = \frac{G}{D}, \quad b = \frac{B}{D} \quad (2)$$

3. 얼굴 영역 분리와 특징점 추출

입력 영상시퀀스에서 얼굴 영역을 찾는 문제는 본 논문에서 제안한 얼굴 트래킹 시스템에서 매우 중요하다. 본 논문에서는 입력된 영상 시퀀스에서 얼굴 영역을 분리하는 방법으로 인접한 두 영상사이의 차영상과 얼굴이 가지는 R, G, B 값들을 동시에 사용한다. 먼저, 동영상에서 두 프레임(frame)사이의 차이를 비교하는 가장 간단한 문제는 식(3)과 같이 양쪽 프레임사이의 픽셀을 직접적으로 비교하는 방법이다.

$$\begin{cases} 1 & \text{if } |F(x, y, t) - F(x, y, t+T)| > \tau \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

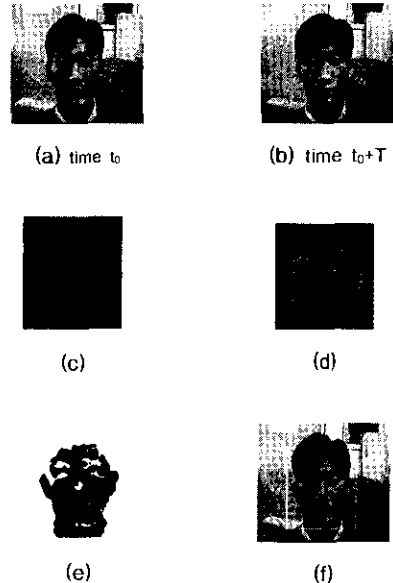


그림 2. 얼굴 영역 분리

그림 2에서 구해낸 얼굴 영역에서 눈이 존재할 가능성이 있는 영역을 직사각형으로 선택한다. 선택한 직사각형 영역의 RGB 영상을 그레이(grey) 영상으로 변환시킨 후 다시 임계값(Threshold T = 190)을 사용하여 이진영상으로 전환한다.

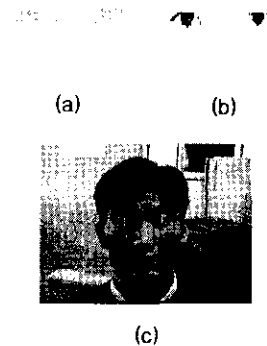


그림 3. 추출된 눈의 위치

이 변화된 직사각형 영역에서 식 (4)와 (5)를 적용하여 본 논문에서 특징점으로 사용하는 입력 영상의 눈의 위치를 추출해 낸다.

$$\bar{x} = \sum_{i=1}^m \sum_{j=1}^n B[i, j] = \sum_{i=1}^m \sum_{j=1}^n iB[i, j]$$

$$\bar{y} = \sum_{i=1}^m \sum_{j=1}^n B[i, j] = \sum_{i=1}^m \sum_{j=1}^n jB[i, j]$$

$$\bar{x} = \frac{\sum_{i=1}^m \sum_{j=1}^n iB[i, j]}{A} \quad (4)$$

$$\bar{y} = \frac{\sum_{i=1}^m \sum_{j=1}^n jB[i, j]}{A} \quad (5)$$

식(4), (5)에서 i 는 x 방향으로의 픽셀, j 는 y 방향으로의 픽셀, A 는 직사각형 영역에서 검은색 픽셀의 총 개수를 나타낸다. 그림 3은 추출된 눈의 위치를 나타내고 있다. 또한, 20대 남자의 표준얼굴에서 구한 양쪽 눈과 입의 이루는 거리비(양쪽 눈 사이의 거리 : 양쪽 눈이 이루는 중점과 입까지의 거리 = 1 : 1.14) 를 적용하여 여기서 구한 양쪽 눈의 위치와 입의 위치 선택을 바탕으로 표준 템플리트를 변형시키게 된다.

4. 변형가능 템플리트와 텍스처 매핑

식(4), (5)에서 구한 눈의 위치 정보와 입의 위치 선택을 가지고 표준 템플리트를 변형가능 템플리트로 변화시켜 입력 영상에 정합된 모습이 그림 4에 나타나 있다.



그림 4. 정합된 템플리트

하지만, 이 변형가능 템플리트는 실제 얼굴의 형태와 방향정보만 가질 뿐이다. 결국, 트래킹하는 변형가능 템플리트와 실제 입력 영상사이의 직접적인 비교가 불가능하다. 따라서, 앞서 말한바와 같이 유일한 데이터인 20대 남자의 표준얼굴을 변형가능 템플리트를 통해 어핀 변환(affine transform)시킨 텍스처 매핑 결과영상과 실제 입력 영상과의 차이를 비교한다. 두 영상사이의 차는 식

(6)을 이용해 구해질 수 있다.

식 (6)에서 $F(x, y, t_0)$ 는 입력 영상을 나타내며, $F_D(x, y, t_0)$ 는 텍스처 매핑결과 영상이다.

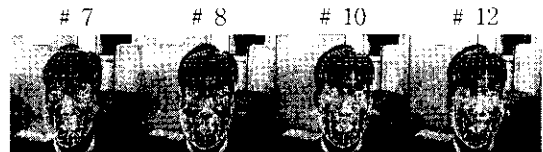
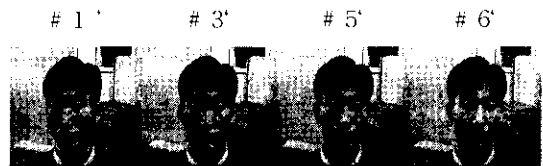
$$error = \frac{|F(x, y, t_0) - F_D(x, y, t_0)|}{N \times M} \quad (6)$$

여기서, N 과 M 은 얼굴이 이루는 각각의 x 방향, y 방향으로 픽셀수를 나타낸다. 그림5의 (a)는 입력 영상이고 (b)는 텍스처 매핑된 결과 영상이다.



(a) (b)
그림 5. 텍스처 매핑 결과

5. 실험 결과



1' # 3' # 5' # 6'
7 # 8 # 10 # 12
그림 6. 얼굴 트래킹 결과 영상

본 논문에서는 CCD 카메라를 사용하여 초당 3

프레임씩 10초간 입력 영상을 획득하였다.

그림 6은 입력된 영상 시퀀스에서 제안된 알고리즘을 통하여 구현된 얼굴 트래킹 시스템의 결과 영상이다

그림 6에서 보면 실제 영상의 첫 번째 프레임에 정합된 템플리트를 사용하여 연속된 프레임에서 얼굴 영역을 매우 근접하게 브래킹(tracking)하고 있음을 확인할 수 있다. 이것을 보다 정확하게 표현하기 위해 식 (6)을 통해 살펴본 얼굴 트래킹의 정확도가 그림 7에 나타나 있다.

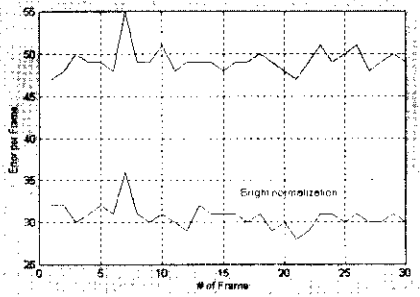


그림 7. 프레임당 트래킹 오차

그림 7에서 보면 텍스처 매핑 결과 영상과 입력 영상과의 휘도치 보정(Brightness normalization)을 통하여 에러가 상당히 줄어들음을 확인할 수 있다.

6. 결론

본 논문에서는 특징점 추출에 기반한 변형가능 템플리트를 이용해 동영상에서의 얼굴 트래킹에 관한 알고리즘을 제안하고 실험을 통하여 그 결과를 확인해 보았다. 실험결과를 토대로 입력된 얼굴의 움직임이 상대적으로 작은 경우 첫 번째 프레임에 정합된 템플리트를 가지고 연속된 영상에서의 얼굴을 매우 정확하게 트래킹함을 볼 수 있었다. 또한, 텍스처 매핑시킨 결과 영상의 휘도치 보정을 통하여 트래킹 오차를 감소시킬 수 있었다.

본 논문에서 사용한 이 변형가능 템플리트는 얼굴 표정의 인식 및 화상 통신, 얼굴 변형 및 합성등 여러 분야에서 응용할 수 있을 것이라 기대된다.

참고 문헌

[1] Chellappa, Wilson, and Sirohey, "Prolog to

Human and Machine Recognition of Faces : A Survey," Proceedings of The IEEE., Vol. 83, No. 5, May 1995, pp.704-740.

[2] Alexandros Eleftheriadis, "Automatic location of faces and facial features in video sequenxes," International Workshop on Automatic Face and Gesture Recognition pp142-147 1995

[3] Henry A. Rowley, Schumect Baluja, and T. Kanade, "Human Face Detection in Visual Scenes," Technical Report CMU-CS-95-158, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 1995.

[4] Roberto Brunelli and Tomaso Poggio , "Face Recognition : Feature versus Templates," IEEE Trans. on PAMI, vol 15 , No 10 , 1993

[5] Manjunath, Shekhar, Chellappa, and Malsburg, "A Robust Method For Detecting Image Features With Application Face Recognition and Motion Correspondence," Proceedings of IEEE, pp.217-220, 1992.

[6] Guangzheng Yang, Thomas S. Huang, "Human Face Detection in a Complex Background," Pattern Recognition , vol 27 , No. 1, pp. 53-63 , 1994

[7] M.A.Sid-Ahmed, "Image Processing: Theory, Algorithms & Architectures," MaGraw-Hill Interna tional Editions, pp.277-375

[8] L. Harmon and W. Hunt, "Automatic Recognition of Human Face Profiles," Computer Graphic and Image Process, Vol. 6, pp.135-156, 1977..

[9] Toshiaki Kondo, Hong Yan, "Automatic Human Face Detection and Recognition under Non-uniform Illumination," The Fourth International Conference on Control, Automation, Robotics and Vision, Singapore,3-6 December, 1996