

MPEG-4 코딩을 위한 눈 추적과 애니메이션

박동희*, 배철수*, 나상동**

*관동대학교 전자통신공학과, **조선대학교 컴퓨터공학과

*e-mail : easthee1@chollian.net

Eye Tracking and synthesize for MPEG-4 Coding

Dong-hee Park*, Cheol-soo Bae*, Sang-Dong Na**

*Dept of Electronics Communication, Kwandong University

**Dept of Computer Engineering, Chosun University

요 약

본 논문에서는 3D 모델의 눈 변형을 계산하기 위해 검출된 눈 형태를 이용한 눈 움직임 합성 방법을 제안하였다. 얼굴 특징들의 정확한 위치 측정과 추적은 MPEG-4 코딩 시스템을 기반으로 한 고품질 모델 개발에 중요하다. 매우 낮은 비트율의 영상회의 응용에서 시간의 경과에 따라 눈과 입술의 움직임을 정확히 추적하기 위해 얼굴 특징들의 정확한 위치 측정과 추적이 필요하다. 이들의 움직임은 코딩되어지고 원격지로 전송되어 질 수 있다. 애니메이션 기술은 얼굴 모델에서 움직임을 합성하는데 이용되어진다. 본 논문에서는 얼굴 특징 검출과 추적 알고리즘으로 잘 알려지고, 효과적으로 향상시킬 수 있는 휴리스틱 방법을 제안하였다. 본 논문에서는 눈 움직임의 검출뿐만 아니라 추적, 모델링에도 초점을 두었다.

1. 서 론

영상 분석에서, 윤곽선과 영역 같은 구조적 특징 또는 대상으로 분류할 수 있다. 이들 영상 특징 또는 대상은 매우 낮은 비트율로 영상을 인코딩 하는데 이용되어지고 있다. 모델 기반 코딩으로 알려진 이러한 연구는 최근 영상 분석과 컴퓨터 그래픽과 관련하여 주목을 받고 있다. 현재 3D 모델 기반 코딩은 인간의 얼굴 영상에 초점을 두고 있다. 모델 기반 코딩에 관한 많은 방법들이 발표되고 있고[3, 9] 얼굴 표정 분석과 합성에서 자동 얼굴 특징 검출과 추적, 정확한 특징 및 움직임을 찾는 것이 여전히 해결해야 할 과제로 남아 있다. 얼굴 특징 검출에는 다양한 방법이 제안되고 있다. 변형 가능한 템플릿 정합[4, 6, 14], 히프 변환, 컬러 영상 처리[2, 13]등이 있다. 변형 템플릿 정합은 템플릿의 에너지 최소화 과정을 거쳐야 위치 최소값을 찾을 수 있기 때문에 정확한 초기 위치 측정이 요구된다. 또 다른 문제는 정합 시 여러 에너지 조건과 가중 인자를 사용하게 됨에 따라 발생하게 된다. 본 논문에서는 이런 문제들을 해결하기 위해서 초기 위치 측정 과정을 향상하였다. 컬러 영상에서 히프 변환과 변형 템플릿 정합이 결합된 것이 매우 좋은 결과를 나타낸다는 것을 알아냈다.

모델 기반 코딩에서 중요한 것은 실제 얼굴영상에서 검출된 모션 파라미터들과 얼굴 모델의 애니메이션 방법을 이용하여 원격지에서 얼굴 움직임과 표정을 합성하는 것이다. 얼굴 표정을 표현하기 위한 여러 접근 방법들로는 특징 검출[1,14], 얼굴 표정분석[11,12], 얼굴 표정합성[7,10]등이 있으며, 눈 표정합성에 관한 논문들도 발표되고 있다. 눈은 얼굴 표정에 가장 확실한 부분이며 눈 움직임의 미세한 변화는 다른 표정을 나타내게 된다. 그러므로 정확한 눈 표정 분석과 합성은 꼭 필요하다.

본 논문에서는 3D 모델의 눈 변형을 계산하기 위해 검출된 눈 형태를 이용한 눈 움직임 합성 방법을 제안하였다. 개별적인 3D 얼굴 모델을 만든 후 얼굴 열의 첫 프레임에 그 모델을 사상하였다. 검출된 눈 특징점을 기본으로 이어지는 프레임들에 눈 움직임을 합성하기 위해서 3D 모델에 변형 파라미터를 사용한다.

본 논문의 순서는 다음과 같다. 2장에서는 정확한 눈 특징점 검출과 추적에 관한 방법을 제시하고 3장에서는 눈 움직임을 표현하는 애니메이션 기법을 제시한다. 4장에서는 실험 결과들을 분석하고 마지막으로 5장에서는 결론을 맺는다.

2. 눈 형태의 검출과 추적

눈을 검출하기 위해 본 논문은 허프 변환과 변형 가능한 템플릿 정합을 사용했으며[4] 정확한 눈을 검출하기 위해서 컬러 정보를 이용하였다. 알고리즘은 다음과 같다.

- 눈에서 두 개의 영역을 결정
- 허프 변환을 이용한 기울기를 사용해 눈의 홍채 검출
- 눈의 경계 검출을 위한 미세 영역 결정
- 눈꺼풀에 대한 초기 근사값을 얻기 위해서 컬러 정보를 이용
- 변형 템플릿을 이용한 눈꺼풀 위치 설정

얼굴 영역을 검출한 후 영상의 왼쪽 상단과 오른쪽 상단의 두 영역은 눈을 검출하기 위해 설정되며 눈 크기에 대한 영역도 또한 얻을 수가 있다. 홍채는 눈의 가장 중요한 형태며 원형을 띄고있기 때문에 원 형태에 대해서는 우선 허프 변환을 이용한 기울기를 사용해서 검출한다[5,8]. 에지의 필수 정보인 위치, 크기, 방향은 Sobel로써 검출되고 최대 크기의 30% 작은 크기로 모든 에지를 제거함으로써 수행을 하게 된다. 그림 1은 허프 변환을 보여주고 있다.



그림 1. 원 추출(영상과 결과, 에지 영상, 허프 공간)
Fig 1. Extraction of Circle(Image and result, edge image, Hough space)

홍채의 윤곽을 제공하지 않는 많은 에지에도 불구하고 허프 공간은 정확한 홍채의 가운데 위치에서 최대값을 보여준다. 원을 검출한 다음, 눈꺼풀의 변형 템플릿은 초기화해야 한다. 모든 파라미터와 모델은 그림2에 나타내었다. 파라미터는 다음과 같다.

$$h_1 = 1.5r_{iris}, h_2 = 0.5r_{iris}, w = 2.2r_{iris} \quad (1)$$

여기서 r_{iris} 는 검출된 원의 라디안이다. 방향 a 는 두 원의 중심점에 의해 결정된다. 초기화된 변형 템플릿은 그림 2에 나타내었다.

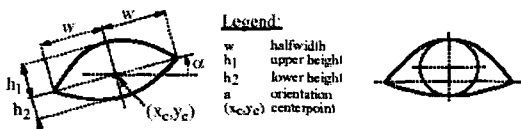


그림 2. 변형 가능한 템플릿(모델, 초기화)

Fig 2. Deformable template(Model, Initialization)

변형 템플릿 정합을 위해 영상 정보의 두 개의 다른 형태가 영역을 생성하는데 사용된다. 전형적인

눈에서 검출된 영상 정보를 그림3에 나타내었다.



그림 3. 에너지 계산을 위해 사용된 영상 영역 (영상, 채도, 에지)

Fig 3. Image regions used for computing the potential energy(Image, saturation, edge)

눈꺼풀의 초기화 단계는 홍채 위치에 변형된 템플릿을 근사시키는 것이며 에너지(E_{sat})를 최소화함으로써 이루어진다.[14]

$$E_{sat} = \frac{-1}{|A_w|} \int_{A_w} \Phi_{sat}(\bar{x}) dA \quad (2)$$

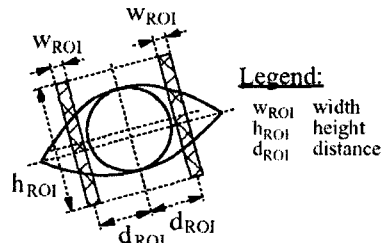
A_w 는 포물선 안쪽 영역이다. 하지만 홍채의 원 안쪽에 있는 것은 아니다. 그리고 $\Phi_{sat}(x)$ 는 컬러 영상의 반전된 채도값이다. 단지 위치만 바뀌기 때문에 이 방법은 수축효과를 가지지 못한다. 다음 단계는 파라미터 h_1 과 h_2 를 근사시킬 것이다. 특히 파라미터 h_1 은 중요한데 눈꺼풀 상단의 큰 움직임 때문이다. 파라미터를 측정하기 위해서 홍채의 양쪽 끝 영역은 다음과 정의된다.

$$w_{ROI} = 5, h_{ROI} = 6r_{iris}, d_{ROI} = r_{iris} + 5 \quad (3)$$

변형 템플릿은 홍채 안쪽 위치에 의존하며 왼쪽이나 오른쪽 영역은 더욱 더 많은 계산이 요구된다. 수평 투영[1]을 사용하고 투영도 안에 두 개의 중요한 반대 기울기를 검출함으로써 눈꺼풀의 상단과 하단 점을 검출할 수 있다.(그림4). 템플릿의 파라미터 h_1 과 h_2 는 다음과 같다.

$$h_1 = h_1 \frac{|y_{up} - y_c|}{h_1^*}, h_2 = h_2 \frac{|y_{low} - y_c|}{h_2^*} \quad (4)$$

y_{up} 과 y_{low} 는 상단과 하단 눈꺼풀 영역 안쪽에서 검출된 점의 y좌표이다. h_1^* 과 h_2^* 는 영역 안쪽의 실제 포물선의 높이이다.



Legend:
wROI width
hROI height
dROI distance



그림 4. 홍채주변의 완전 투영
Fig 4. Integral projections beside iris

마지막 단계는 에너지 (E_{edge})를 최소화함으로써 눈꺼풀에 변형 템플릿을 정합하는 것이다.

$$E_{edge} = - \frac{1}{|B_w|} \int_{B_w} \Phi_{edge}(x) ds$$

B_w 는 포물선의 경계이고 $\Phi_{edge}(x)$ 는 에지 크기이다. 변형 템플릿의 모든 파라미터(위치, 방향, 높이, 폭)는 최소화하는 동안 변화되어진다.

눈 특징의 추적은 눈 특징 검출과 비슷하며 차이점은 다음과 같다.

- 영역뿐만 아니라 홍채 검출을 위한 허프 공간의 크기는 이전 프레임에서 검출된 눈 크기와 위치를 이용함으로써 제한될 수 있다.
- 그림2의 변형 템플릿을 이용하는 것 대신에 이전 프레임의 매치된 템플릿은 새로운 템플릿을 초기화하기 위해서 이용된다.

3. 눈 움직임 애니메이션

눈 검출과 추적 알고리즘은 영상 열의 각 프레임에서 홍채와 눈꺼풀의 윤곽을 검출한다. 이들 눈 모양은 3D얼굴모델에서 눈의 실제 움직임을 합성하는데 이용되며 순서는 다음과 같다.

- 3D 눈 모델을 기반으로 11개의 특징점들은 그림 5에서 보여주며 눈꺼풀 윤곽과 홍채는 영상 열에서 검출된다. 템플릿의 특징점은 포물선 경계의 점들을 계산하고 원의 중심을 이용함으로써 결정될 수 있다.

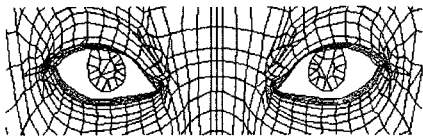


그림 5. 정의된 눈 특징점
Fig 5. Defined eye feature vertices

- 홍채 애니메이션: 영상 열에서 홍채가 검출된 후, 눈 모델의 홍채 중심은 눈 영상의 홍채 중심과 정합 되어진다. 홍채 모델의 크기는 홍채템플릿의 원에서 조정될 수 있다.
- 눈꺼풀 애니메이션: 영상 열에서 눈꺼풀이 검출된 후, 눈 모델의 특징점은 눈 영상의 특징 점들에 의해 고정된다. 특징점과 특징점의 교차는 움직임 벡터(FMV: Feature Motion

Vectors)의 열들로 구성된다. 눈 움직임을 변형하고 애니메이션 하기 위해 비특징점의 움직임 벡터열을 찾아야 하며, 비특징점 주변에 있는 형태 움직임 벡터를 보간법을 사용하여 유도할 수가 있다. 보간법 절차에서 비특징점의 새로이 얻어지는 움직임 벡터는 FMV 열속에 넣는다. FMV의 크기는 동적으로 증가하게 된다. 눈 모델(v로 표기)에서 비특징점의 움직임 벡터는 다음과 같다.

$$v = \sum_{n=1}^N (w(d_n) \cdot v_n) \quad (6)$$

여기서 d_n 은 비특징점(v)간의 거리이고 특점 $v_n, n=1, 2, \dots, N; (d_1, \dots, d_N)$ 은 증가되는 차수에서 재배열되고 $w(d_n)$ 은 작은 입력 d_n 에 큰 출력 값을 가지는 가중치 함수이다.

$$w(d_n) = \frac{d_n}{\sum_{n=1}^N d_n} \text{ where } n' = (N+1) - n \quad (7)$$

이 정점들을 근간으로, 변형된 눈 모델은 각 프레임에서 만들어질 수 있다. 합성된 눈 움직임 결과는 그림 6에서 보여주고 있다.

- 최종적으로 합성된 눈 모델은 실제 눈 표정을 합성하기 위해서 원영상(첫 프레임)에 의해 텍스처 사상될 수 있다.

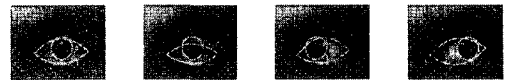
4. 실험 결과

제안된 알고리즘을 확인하기 위해 3118조각의 고해상도 얼굴 모델을 사용하였고, 다양한 눈 모델들의 컬러 영상을 가지고 실험을 하였다. 모의 실험에서 허프 변환은 잡음뿐만 아니라 홍채의 윤곽에 의해서 만들어지지 않는 에지에도 매우 강하다.

홍채의 동작 및 추적 결과, 하나의 영상 열에서 눈꺼풀의 움직임을 그림 6에 나타내었다.



(a) 원영상



(b) 추출된 눈 템플릿



(c) 정합되고 동작되는 3D 모델
그림 6. 눈 움직임을 보여주는 열

Fig 6. Sequence showing eye movement

눈 움직임은 다음과 같이 합성되어진다. 첫째 눈 움직임은 개별화된 와이어 프레임 모델의 변형으로 모델화 된다.(그림7의 윗그림). 이때 연속적인 프레임에서 와이어 프레임 모델은 열의 첫 프레임을 사용함으로써 텍스처 사상되어질 수 있다.(그림7의 아래그림)

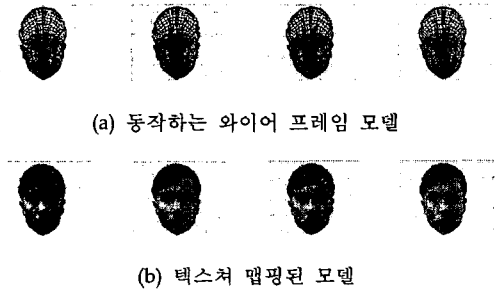


그림 7. 개별적 모델과 원영상, 변형된 모델과 합성된 영상

Fig 7. Individual model and original image, deformed model and synthesized images

본 논문의 모의 실험 결과로부터 제안된 알고리즘이 눈 움직임에 대해 잘 동작함을 볼 수가 있었으나 모델링, 분석, 합성의 시간적 복잡성은 결과적으로 증가하였다.

5. 결론

본 논문에서는 저 비트율 코딩 모델 기반에서 눈 움직임 검출과 추적, 눈 움직임 애니메이션 방법을 제안하고 있다.

히프 변환은 잡음뿐만 아니라 홍채의 윤곽에 의해서 만들어지지 않는 에지에도 매우 강함을 알 수 있었고, 눈 움직임에 대해 잘 동작함을 볼 수 있었다.

향후 과제로는 모델링, 분석, 합성의 시간적 복잡성을 줄여야 할 것이고 입술의 움직임과 네트워크 상에서 실시간 시스템의 연구가 진행되어야 할 것이다.

참고문헌

[1] R. Brunelli and T. Poggio. Face Recognition: Features versus Templates. IEEE Transactions on Pattern Analysis and Machine Intelligence. 15(10):1042-1052, Oct. 1993.

[2] T. Chang and T. Huang. Facial Feature Extraction from Color Images. In Proc. 12th IAPR Int. Conf. on Pattern Recognition, Jerusalem, Israel. volume II. pages 39-43. IEEE Computer Society Press. October 9-13 1994.

[3] C. Choi and T. Takebe. Analysis and Synthesis of Facial Image Sequences in Model-Based

Image Coding. IEEE Transactions on Video Technology. 4:257-275. June 1994.

[4] G. Chow and X. Li. Towards a System for Automatic Facial Feature Detection. Pattern Recognition. 26(12):1739-1755, Dec. 1993.

[5] E. davies. A Modified Hough Scheme for General Circle Location. Pattern Recognition. 7:37-43, Jan. 1988.

[6] J. Dong and F. Lai. Region-Based Template Deformation and Masking for Eye-Feature Extraction and Description Pattern Recognition. 30(3):403-419, Mar. 1997.

[7] P. Ekman and W. Friesen. Facial Action Coding System. New York: Consulting Psychologists Press. 1977.

[8] P. Kierkegaard. A Method for Detection of Circular Arcs Based on the Hough Transform. Machine Vision and Applications. 5:249-263, 1992.

[9] D. Pearson. Developments in Model-Based Video Coding Proc. of the IEEE, 83(6):892-906, June 1995.

[10] S. Platt and N. Badler. Animating Facial Expressions, Computer Graphics. 13(3):245-252, 1981.

[11] L. Tang and T. Huang. Quantifying Facial Expressions: Smiles, In Proc. of the Int, Workshop on Coding Techniques for Very Low bit-rate Video. pages 345-350, 1994.

[12] D. Terzopoulos and K. Waters. Analysis and Synthesis of Facial Image Sequences Using Physical and Anatomical Models. IEEE Transactions on Pattern Analysis and Machine Intelligence. 15(6):569-579, June 1993.

[13] H. Wu. T. Yokoyama, D. Pramadihanto, and M. Yachida. Face and Facial Feature Extraction from Color Image. In Proc. 2nd IEEE Int. Conf. on Automatic Face and Gesture Recognition, Killington, Vermont, USA. pages 345-350, Oct 14-16 1996.

[14] A. Yuille, P. Hallinan, and D. Cohen. Feature Extraction from Faces Using Deformable Templates. International Journal of Computer Vision. 8(2):99-111, 1992.