비전 기반 3차원 얼굴 모델의 실시간 표정 제어

김정기[O] 민경필[1] 전준철[2]
경기대학교 정보과학부
conel20[O]@hanmail.net, {cabbi[1], jcchun[2]}@kyonggi.ac.kr

Real-time Expression Control of Vision Based 3 Dimensional Face Model

Jung Gee Kim[0] Kyong Pil Min[1] Junchul Chun [2]
Dept of Information Science, Kyonggi University[O12]

요 약

본 논문은 연속적으로 입력되는 2차원 얼굴 영상에서 얼굴의 특징 영역들을 추출하여 3차원 얼굴 모델의 표정을 실시간으로 제어하는 방법에 관한 연구이다. 2차원 얼굴 영상에서 얼굴을 추출하기 위해 Hue, Saturation 색상 값을 사용하며, 두 가지 색상 값을 이용하여 피부색과 배경색을 분리 함으로써 얼굴 영역을 추출 할 수 있다. 추출 된 얼굴에서 특징 영역인 눈 , 코, 입술 영역 등의 위치를 각각의 영역에 적합한 추출 방법을 이용하여 추출한 뒤, 프레임 별로 영역들의 움직임을 비교함으로써 영역의 움직임 정보를 획득 할 수 있다. 이 정보를 3차원 얼굴 모델에 적용하여 2차원 동영상에서 획득된 대상의 얼굴의 표정을 3차원 얼굴 모델에 실시간으로 표현 할 수 있도록 한다.

1. 서 론

2차원 영상인 사진이나 동영상에서 사람의 얼굴을 검출하거나 얼굴의 특징영역인 눈, 코, 입 등을 검출 또는 얼굴의 표정을 인식하는 연구가 많이 진행 되어 왔다. 특히 최근에는 단순한 얼굴 검출이나 인식과 같은 2차원 영상 분석뿐만 아니라, 인간과 컴퓨터 간의 상호작용 (Human and Computer Interaction)등 여러 분야에서 그 관심이 증가하고 있다. 사람의 얼굴은 상호간의 있어 매우 중요한 의미를 쉽게 나타나게 된다. 얼굴의 표정으로부터 감정을 전달 받기도 하고, 음성이나 상대방의 입술 모양을 통해서 정보를 전달하기도 한다. 이러한 관심 속에 2차원 영상으로부터 얻은 정보를 3차원 얼굴 모델에 적용하는 방법에 관한 연구들이 많이 진행되고 있다.

2차원 정보를 통해 3차원 얼굴 모델을 제어하는 방법은 크게 세 가지 분류 방법에 의해서 구분될 수 있다. 첫 째, 제 3자 즉, 애니메이터의 간섭을 필요로 하는지 안하는지의 여부에 따라서 제어 방법이 나뉘어질 수 있으며, 둘 째, 얼굴의 특징을 쉽고 정확히 검출하기 위해서 얼굴에 직접 마커(marker)를 부착하는지의 여부에 따라 나눌 수 있다. 마지막으로, 2차원 영상으로부터 정보를 추출하여 3차원 모델을 제어하는데까지 필요한 모든 처리가 실시간으로 가능한지의 여부에 따라 나눌 수 있게된다.

일반적으로 3차원 작업은 애니메이터의 간섭의 의해서 모델의 생성이나 변형이 이루어 진다. 모든 작업이 수작업이기 때문에 숙련된 작업자에 의해서 정교한 제어 효과를 얻을 수 있다. 하지만, 많은 시간이 소비된다는 단점을 가지고 있다.

얼굴 영상에서 각각의 특징 영역들을 추출 할 때는 많은 제약들이 따르며, 그 조건에 맞지 않을 경우 정확한 검출을 하지 못하게 된다. 그런 어려움을 제거하기 위해 사람의 얼굴에 직접 마커를 부착하여 특징 영역 및 표정 변화 시 발생하는 근육의 움직임을 추출할 수 있다. 정확한 결과를 추출할 수 있지만, 얼굴에 마커를 부착한다는 불편성과 마커의 움직임을 추적하는 장비들의 항상 필요하다는 단점이 존재한다.

특징 영역들의 추출을 정확하기 위해서 과다한 연산을 하거나, 특징 영역들을 마커와 같이 영상에서 직접 정의 또는 3차원 얼굴 모델의 사실성을 증대시키기 위해서 많은 보간 연산을 할 경우 실시간 처리가 불가능하게 된다. 실시간 처리가 가능하도록 각 단계의 연산을 줄이도록 하는 연구들이 제안되고 있다.

본 논문에서는 3차원 모델을 제어하기 위해서 사용자의 간섭없이 모든 과정을 실시간으로 처리할 수 있는 방법을 제안하도록 한다. 2차원 동영상에서 사람의 얼굴을 검출하기 위한 방법 중 가장 빠른 속도를 낼수 있는 색상 정보를 이용하였고, 얼굴의 특징 영역들을 추출하기 위해 각 영역에 특성에 맞는 최소 연산 방법을 적용한 후, 각 특징 영역들의 움직임을 각 프레임 별로 추출하였다. 이를 3차원 얼굴 모델에 적용함으로써 입력 되는 대상의 얼굴의 표정을 3차원 얼굴 모델에 나타내도록 하였다.

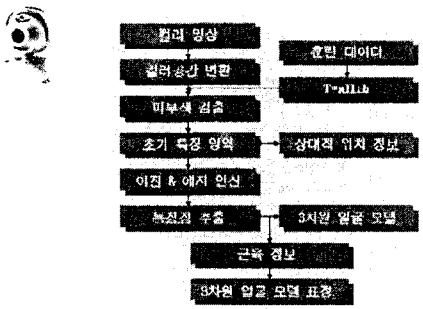
이후 논문의 본론에서는 얼굴 검출에 이용되는 색상 정보와 검출 방법을 설명하고, 각 특징 영역에 적합한 추출 방법, 그리고 3차원 얼굴 모델에 각 특징 영역들의 움직임을 적용하는 방법을 설명 하도록 한다. 마지막으로 결론과 향후 연구 방향에 대해 제시하도록 한다.

2. 본 론

본 논문은 동영상으로부터 얼굴 영역 및 특징을 추출한 후, 3차원 모델을 제어하는 방법을 사용자의 간섭없이 실시간으로 처리하는데 목적이 있다. 제안 방법은 [그림 1]과 같이 간략하게 설명할 수 있다.

입력 데이터로는 연속적으로 입력되는 컬러 영상을 사용하며, 일반적인 색상 공간인 RGB 색상 공간에는 색상 정보와 밝기 정보가 함께 섞여있기 때문에, HT(Hue-Tint) 색상

공간으로 변환하여 얼굴검출에 필요한 색상 정보만을 추출한다. HT공간 내에서 피부색들의 분포도의 특징을 이용하여 피부 영역을 검출 하도록 하고, 불필요한 잡음을 제거하기 위하여 median filter를 적용한다. 검출된 피부 영역에서 얼굴에 대한 특징 영역의 구조적 정보와 히스토그램 정보를 이용하여 눈과 입에 대한 특징 영역의 위치를 추출 하며, 좀더 정확한 특징 영역의 위치를 찾기 위해 추가적인 계산으로 에지 연산과 이진화 과정을 추가하였다. 추출된 특징 영역들의 움직임을 근육 정보와 함께 3차원 얼굴 모델에 적용함으로써 3차원 얼굴 모델의 표정을 제어 할 수 있다.

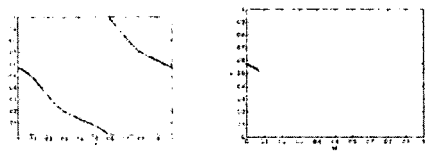


[그림 1] 시스템 개략도

2.1 색상 정보를 이용한 얼굴 검출

얼굴 검출 기법으로 이해하기 쉽고, 비교적 정확한 검출 결과와 빠른 연산을 보장하는 방법으로 컬러 영상으로부터 피부 색상 정보를 이용하는 방법이 있다. 색상 정보는 영상으로부터 피부를 인지할 수 있는 가장 효율적인 방법이며, 빠른 처리 능력과 얼굴 형태의 기하학적 변화에 많은 영향을 받지 않는 특징을 갖고 있다. 일반적인 문제점으로는 피부와 유사한 색으로 인하여 얼굴을 검출하지 못한다는 점과 밝기 정도에 따른 피부색 오검출 문제가 있다. 이러한 문제점들을 해결하기 위해 RGB 색상 공간을 YCbCr, HSV, 그리고 TSL과 같은 다양한 색상 공간으로 변환하여 검출하는 방법에 관한 연구가 많이 진행되고 있다.[3][4]

색상 정보를 이용한 검출을 하기 위해서는 먼저, 피부 색상과 유사색을 쉽게 구분지을 수 있으며, 피부 색상을 수식으로 쉽게 표현할 수 있는 색상 공간을 찾아야 한다. 그러한 색상 공간을 찾기 위해서 색상 값에 대한 분포도를 이용한다. [그림 2]에서 볼 수 있듯이 HT 조합에 의한 색상 공간에서의 피부색 분포가 다른 분포에 비해 비교적 작은 영역에 밀집되어 있으며, 직선 형태로 색상 정보가 표현되기 때문에 쉽게 수식화 할 수 있다는 장점이 있다. 또한, HT 조합에 의한 색상 공간은 밝기에 대한 영향을 덜 받기 위해서 H(Hue) 색상 정보를 포함하고 있다. H값은 붉은색, 초록색, 노란색, 자주 빛과 같은 현저한 색을 정의하는데 이용되며, 급격한 빛의 변화가 아닌 상태의 빛의 변화에도 H값의 수치적 변화가 생기지 않는다는 특성으로 인해 빛에 대한 영향을 덜 받을 수 있는 장점이 있기 때문에, 본 연구에서는 HT 색상 정보를 이용하도록 한다.



(a) 일반영상 (b) 피부색 영상

[그림 2] HT 컬러 모델의 분포도

HT 색상 공간에서 피부 영역을 검출하기 위해서는 피부색상과 다른 색상들을 정확히 구분할 수 있어야 한다. HT 색상 공간에 나타나는 색상의 분포는 [그림 2]에서 볼 수 있듯이 직선 형태로 나타난다. 그러므로, 직선의 방정식을 이용하여 색상을 표현하고, 그 방정식을 (식 1)과 같이 피부 색상이 나타나는 영역으로 제한하도록한다.

$$f(h)=\frac{T_{\min}-T_{\max}}{H_{\max}-H_{\min}}(h-H_{\min})+T_{\max} \quad ,H_{\min}\le h\le H_{\max} \tag{1}$$

$$d(x,y)=\left|f(x)-c(x,y)\right| \tag{2}$$

(식 1) 은 H와 T의 최대값과 최소값에 해당되는 두 개의 점을 지나는 직선의 방정식을 나타내며, (식 2)는 입력 영상의 각 픽셀의 색상 값을 HT로 변환 한 값과 직선간의 거리를 측정하는 수식으로 피부 색상인지 아닌지를 판별하게 된다. [그림 3]은 제안 방법을 이용하여 2차원 일반 영상과 동영상으로부터 얼굴 영역을 검출한 결과이다.

(a)원영상 (b) 피부 검출 영상

[그림 3] 피부 영역 검출 결과

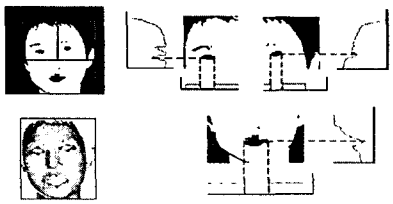
2.2 특징 영역 검출

제안 방법에의해 검출된 결과는 외부 조명이나 잡음으로 인해 피부 영역이 아닌 미세한 부분도 검출 될 수 있다. 이러한 잡음을 제거를 위한 후처리 과정으로 미디언 필터를 적용한다.

검출된 피부 영역에서 얼굴 영역만을 선택하기 위해서 검출된 특징들의 구조를 분석한다. 일반적으로 얼굴에서 눈, 눈썹, 입 영역은 얼굴의 다른 부분에 비해서 낮은 밝기 값을 갖는다. 그리고 눈과 입은 수평 방향으로 위치하기 때문에, 영상의 수평 성분을 검색하면 쉽게 눈과 입 영역을 찾을 수 있다.

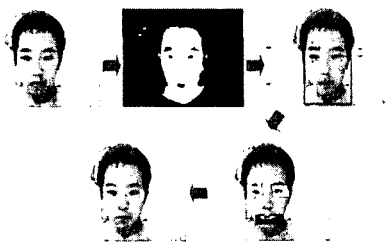
정확한 검출을 위해 얼굴 영역을 상·하·좌·우로 4등분으로 나누어서 얼굴의 특징 영역들이 검출될 수 있는 후보 영역을 설정한다. 윗 부분에 나뉘어진 두 개의 영역에서는 눈이 검출 될 수 있으며, 아랫 부분으로 나뉘어진 영역에서는 입이 검출 될 수 있다. 각 영역에 대해 먼저 각 열들의 합에 대한 누적 그래프를 만들어 보면, 그래

프에서 발생하는 최소값을 검출할 수 있다. 검출된 최소값에 대한 행의 값을 그리는 그래프를 작성하면, 그 그래프에서의 최소값을 발견할 수 있다. 이러한 최소값들의 위치 정보를 분석하면, [그림 4]와 같이 얼굴 영역에서 눈과 입 영역에 대한 위치를 검출할 수 있다.



[그림 4] 특징 영역 검출 결과

사람의 얼굴 구조는 눈과 입의 위치 정보를 기준으로 특정 비율을 갖고 있으므로, 검출된 눈과 입의 위치를 기준으로 다른 특징 후보 영역이 존재할 수 있는 블록들을 설정하면 보다 정확한 검출이 가능한다. 특징 영역들 중 눈이나 눈썹의 같이 주변의 색인 피부색과의 구별이 명확한 경우, 이진화 연산을 통해 피부색과 검은색의 명암도 차이를 이용하여 눈과 눈썹을 얼굴에서 분리 할수 있다. 입술 영역의 경우 피부색 검출시 얼굴 영역에서 검출되지 않은 부분의 위치를 추정하여 얻을수 있다. 하지만 조명의 영향이나 화장으로 인해 입술의 색깔이 변형 되었을 경우 피부색으로 인지하여 정확한 특징점을 추출할수 없는 경우가 발생하게 된다. 이 문제를 해결하기 위해 에지 정보를 추가 한다. 에지 연산으로는 많이 사용되고 있는 3X3 소벨 마스크를 입술 영역 블록에만 적용하였다. [그림 5]는 제안 방법에 의해 얼굴의 특징 영역들이 검출되는 과정을 순서대로 보여준다.



[그림 5] 특징점 추출 처리 과정

2.3 3차원 얼굴 모델의 제어

2차원 동영상으로부터 특징점을 추출한 후, 그 영역들의 움직임을 직접 모델에 적용 했을 경우 얼굴에서 움직이는 모든 부분 대한 움직임을 추출한 것이 아니기 때문에 3차원 모델에 왜곡 현상이 발생한다. 이로인해 자연스러운 표정을 나타내지 못하게 된다. 이러한 점을 보안하기 위해 사람의 표정 변화시에 발생하는 근육들의 움직임을 모델화하여 3차원 얼굴 모델에 적용하고 있다. 근육 모델에는 Free form deformaiton, Spline psuedo muscles, Layered spring meshes, Vector Model 등이 있다.

본 논문에서는 Vector Model 중에서 계산량이 적은 Waters의 선형 근육모델을 이용하여 근육의 움직임 정보를 추가하였다.[5]

[그림 6] 3차원 얼굴 모델의 제어

[그림 6]은 추출된 28개의 특징 영역과 Warters가 제안한 선형 근육들을 표시한 것이다. 각각의 제어점들이 움직일 경우, 제어점이 속한 근육의 영향으로 인해 그 주변 점들이 움직이게 된다. 근육의 영향권에 속하는 주위의 점들은 제어점이 움직인 방향으로 해당 근육과의 거리에 반비례하여 이동하게 된다. 단순히 제어점만을 움직였을때 보다 좀더 사실성 있는 표정을 생성할수 있게 된다.

3. 결론 및 향후 연구 과제

본 논문에서는 실시간으로 2차원 동영상으로부터 검출된 값을 이용하여 3차원 얼굴 모델에 입력하여 제어하는 방법을 제시하였다. 실시간 작업을 가능하게 하기 위하여 계산량이 적은 색상 정보를 이용하여 얼굴을 검출하고, 특징 영역의 특성에 적합하며 계산량이 적은 이진 연산과 에지 연산을 이용하여 특징 영역을 검출한다. 그리고 선형 근육 모델을 통한 얼굴 모델의 제어 방법을 제안한다.

향후 연구 과제로는 환경 변화에 민감하지 않으며, 다양한 표정 제어가 가능해야할 것이다. 피부색 검출시에 오류가 발생하지 않도록 하는 제약 조건인 과도한 조명, 카메라와 얼굴의 각도, 배경의 복잡성등의 조건들에 대한 제약들의 영향력을 줄일수 있는 방법을 고려해야 한다. 또, 다양한 표정 제어를 위해서, 더 많은 특징점들의 추출과 직선 방향에 대한 영향을 주는 선형 근육 모델의 보완책에 대한 연구가 필요하다.

4. 참고 문헌

[1] Frederic Pighin, Jamie Hecker, Dani Lischinski, Richard Szeliski, and David Salesin, "Synthesizing Realistic Facial Expressions from Photographs." SIGGRAPH '98, 1998.

[2] I.A.Essa, S.Basu, T.Darrell, A. Pentland, Modelin g, Tracking and Interactive Animation of Faces and Heads using Input from Video, Proceedings of Computer Animation June 1996 Conference.

[3] R. L. Hsu, M. A. Mottaleb, and A. K. Jain. Face detectionin color images. IEEE Trans. Pattern Analysis and Machine Intell., 24:696-706, 2002.

[4] J.C. Terrillon, M. N. Shirazi, H. Fukamachi, and S. Akamatsu, "Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images," Proc. IEEE Int'l Conf. on Face and Gesture Recognition, pp. 54-61, 2000

[5] F. I. Parke, Keith Waters, "Computer Facial Animation", A K Peters Wellesley, 1996