

가변템플릿과 신경회로망을 이용한 실시간 눈 영역의 추적

김도형^U 이선화 이학만 차의영
부산대학교 일반대학원 전자계산학과

(dhkim, somy77, berserk, eycha)@harmony.cs.pusan.ac.kr

A Realtime Tracking of Eye Region Using Deformable Template and Neural Network

Do-Hyung Kim^U Seon-Hwa Lee Hack-Man Lee Eui-Young Cha
Dept. of Computer Science, Pusan National University

요 약

본 논문에서는 다양한 배경을 가지는 연속적인 얼굴 영상에서 실시간으로 눈의 위치를 자동적으로 추출하는 방법에 대하여 제시한다. 얼굴 요소 중에서 눈은 얼굴 인식 분야에 있어서 중요한 특징을 나타내는 주 요소로써 주로 히스토그램 분석과 색상 정보를 이용하여 눈 영역의 윤곽을 추출하는 방법이 제기되고 있다. 본 논문에서는 명암의 변화에도 비교적 적응력이 강한 이진화 기법을 사용하여 원 영상을 이진화하고, 가변 템플릿(Deformable Template)방법을 사용하여 후보 영역을 추출한다. 이러한 후보영역들은 ART2 신경회로망을 이용하여 병합되며, 병합된 후보 영역들은 얼굴 요소의 기하학적 사전지식을 기반으로 검증되어, 시간에 따라 모양변화가 급변하는 눈 영역에 대한 실시간 추출을 가능하게 한다. 이상의 연구 결과는 교통사고 방지를 위한 눈의 졸림감지 등의 응용 시스템에 이용될 수 있다.

1. 서 론

움직이는 물체를 자동으로 추적하고, 정보를 획득하여 처리하는 문제는 영상 처리기술의 발달과 더불어 컴퓨터 비전, 패턴인식, 신경망 같은 다양한 분야에 걸쳐 활발히 연구되고 있으며 많은 응용 분야를 가지고 있다[1,2,3]. 특히 얼굴 요소의 자동 추출에 관한 연구는 각 얼굴 요소의 특징을 추출한 정보를 이용하여, 얼굴인식 및 표정 인식 분야에서도 적용될 수 있다. 현재까지 얼굴 요소 추출에 대한 연구는 정지영상을 대상으로 한 연구들이 주를 이루고 있으나, 실제에서의 응용을 위해서는 동영상상 실시간으로 처리되어야 하며, 이를 위해서는 효과적인 처리시간의 감소와 그 정확성이 동시에 고려되어야만 한다.

본 논문에서는 다양한 배경과 명암을 가지는 동영상에서, 얼굴 요소 중에서도 중요한 정보를 가지고 있는 눈 요소의 실시간 추출에 관한 방법을 제시한다. 추출하고자 하는 눈 요소는 눈의 깜박임등과 같이 대상의 모양변화가 매우 다양하고 순간적인 비강체이기 때문에 그 처리가 매우 까다롭다. 이에 대상 추출 영역의 변화에 따른 실시간 가변 템플릿(Deformable Template)방법과 ART2 신경회로망, 그리고 눈 요소의 기하학적 고찰에 따른 사전지식을

병행하여 그 정확성을 높이는 방법을 제안한다. 또한 실시간으로 추출하기 위한 효과적인 처리속도 감소방법에 대해서도 설명한다.

본 논문의 구성은 다음과 같다.

2장에서는 움직이는 물체로부터 이동 정보를 검출하는 추적 방법에 관한 관련연구들에 대하여 설명하고, 3장에서는 다양한 배경과 명암을 가지는 정지영상에서 눈의 위치를 결정하는 알고리즘을 제안한다. 4장에서는 실시간으로 눈의 이동을 추적하는 방법을 제시하고, 5장에서는 실험 환경 및 결과를 분석하고, 6장에서는 결론을 맺는다.

2. 추적 관련 연구

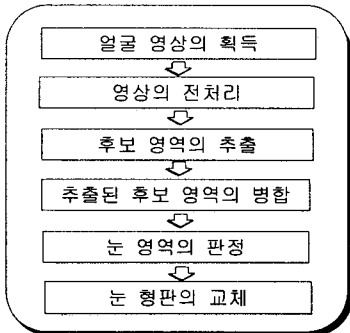
움직이는 물체로부터 이동정보를 검출하는 추적 방법으로는 크게 중심점 추적 방법과 상관추적 방법이 있다. 중심점 추적 방법은 각 화면에서 물체 자체 또는 특징부분의 무게중심을 추출하여 그 무게중심의 변위로부터 이동 정보를 계산하는 방법으로써 대상 물체 주위의 잡영 및 환경에 민감한 단점이 있다[4]. 따라서 중심점 추적 방법은 실제 적용을 위해서는 많은 문제점을 내포하고 있기 때문에 동영상의 시간적, 공간적 상관관계를 이용하여 움직임 벡터를 추정하고 추정된 움직임 벡터를 표적의 추적 방법에 활용하는 상관 추적 방법이 많이 연구되고 있다. 대표적인 상관 추적 방법으로는 화소 재귀적 방법, 블록

[†] 본 논문은 정보통신연구진흥원 대학 S/W연구센터 지원사업비 지원으로 수행되었음.

정합 방법, 그리고 화소 정합 기법 등이 있다. 화소 계극적 기법은 화소 단위로 움직임 벡터를 추정하여 이동을 추적하는 기법으로서 복잡한 움직임 추적에 유용하지만 표적의 경계부분 움직임의 추적에 한계가 있다. 블록 정합 기법은 대상화면에서 하나의 블록을 가지고 예상 변위를 이동하면서 유사성이 가장 큰 변위를 찾아내어 이동 변위를 계산하는 방법으로서, 나누어진 블록 단위로 움직임을 추적하므로 단순한 움직임의 추적에는 적용되나, 복잡한 움직임이나 순간적인 움직임의 변화에는 둔감한 단점이 있다. 화소정합기법은 앞에서 언급한 두 기법을 혼합한 형태로서 그 추적의 정확성은 매우 높은 반면, 계산량 증가에 따른 처리시간의 증가로 실시간 추적에 부적합하다. 이상의 이동 물체 추적 방법들은 실제 많은 응용분야에서 사용되어 우수한 성능을 보이기는 하나, 그 대상 물체가 시간에 따른 모양변화가 없는 강체(rigid body)이다. 따라서 눈과 같이 시간에 따른 모양변화가 급변하는 물체에는 부적합하다.

3. 눈의 위치 결정

연속 얼굴 영상에서의 눈의 위치를 추적하기 위하여 사용되어지는 알고리즘의 전체 흐름도는 [그림 1]과 같다. 먼저 저가의 내장형 캡처 보드와 연결된 카메라로부터 320×240크기의 8bit gray scale 영상을 획득하여, 눈 영역의 특징을 강화하기 위한 전처리 과정을 수행한다. 전처리된 영상에서 눈 형판을 이용하여 후보 영역을 추출한 후, ART2 신경회로망을 사용하여 후보 영역을 병합하고 기하학적 눈의 위치관계를 고려하여 최종 눈 영역을 판정하는 일련의 눈 위치 추적 과정을 거친다. 추출된 한쌍의 눈 영역은 다음 프레임에서의 후보 영역 추출을 위한 형판으로 설정된다.



[그림 1] 전체 알고리즘 흐름도

3.1 전처리

정확한 눈의 위치 결정과 처리시간의 감소를 위해서는 눈 영역을 잘 부각할 수 있도록 이진화를 하는 전처리 과정이 필요하다. 실세계에서 획득되는 얼굴 영상은 명암의 변화에 따라 다양한 밝기값을 가지고, 또한 매순간 형태가 변하므로 이러한 환경에 잘 적용할 수 있는 전처리 방법이 요구된다. 이진화를 하기 위한 임계치 선택기법으로는 크게 단일 임계치를 사용하는 방법과 다중 임계치를 사용하는 방법이 있다. 단일 임계치를 이용하는 방법에는 물체의 영역비에 따라 임계치를 정하는 P-Tile 방법, 마루(peak)와 골(valley)을 이용하는 모드(mode)법, 그리고 반복적 임계치 선택법(Iterative Threshold Selection)등의 방법이 있다. 다중 임계치 방법으로는 Block 이진화 방법과

적응성 임계치 방법(Adaptive Thresholding)등이 있다 [5,6,7]. 하지만 이러한 방법들의 단일 적용으로는 다양한 명암변화에서 눈 영역만을 부각시키는 힘들다. 따라서, 본 논문에서는 얼굴의 명암 변화를 최소화하기 위해서 히스토그램 평활화(Histogram Equalization)방법을 사용하여 원영상을 평활화하고, Sharpening Filter를 사용하여 눈 영역의 모양 및 경계선을 강화한다. 일련의 과정을 거친후 반복 임계치 설정방법을 사용하여 임계치를 설정하면 명암의 변화에도 잘 적용할 수 있는 임계치가 설정되며 이러한 임계치는 대부분의 경우 일정한 값으로 수렴하게 된다. 그러나 설정된 임계치는 전체 영역을 대상으로 하므로 설정된 임계치에서 경험치 α 만큼의 화소값을 제한 임계치가 실제 원영상의 이진화를 위한 임계치로 설정된다.



(a) 입력 얼굴 영상 (b) 이진화된 얼굴 영상
[그림 2] 얼굴영상의 전처리

3.2 후보영역의 추출

전처리 된 얼굴 영상에서 눈 영역일 가능성이 있는 후보 영역은 형판에 의한 유사도 비교방법을 사용하여 추출하게 된다. 첫 프레임에서 유사도를 측정하기 위해 사용된 눈 형판은 전체 영상 크기와 얼굴영상크기를 기반으로 평균적인 눈의 크기를 고려하여 30×15 크기의 이진 형판으로 구성된다. 제작된 형판을 사용하여 원영상의 좌상단점을 기준으로 우하단 방향으로 raster 스캔하면서 해당 블록과 형판의 유사도를 측정하여 설정 임계치 이상의 유사도를 가지는 블록을 눈 후보 영역으로 설정한다. 유사도는 형판과의 Matching Value값을 계산하여 측정하게 되며 눈 영역으로 추정되는 영역의 상좌좌우에서 이러한 후보영역이 집중하게 되는 현상이 발생한다. 이때 처리속도의 향상을 위해 X, Y 방향으로 두 화소씩 스캔한다.



[그림 3] 추출된 눈 후보 영역

3.3 추출 후보영역의 병합

유사도 측정으로 추출된 눈 후보 영역들은 [그림 3]에서와 같이 눈 영역일 가능성이 높은 부분에 집중적으로 나타난다. 집중된 후보 영역부에서 그 무게중심에 위치하는 후보영역이 제일 유사도가 큰 영역이다. 따라서 이러한 무게중심 후보영역을 검출하기 위해 ART2 신경회로망을 사용하여 후보 영역들을 병합하는 과정을 거친다[8]. ART2를 통해 병합된 각각의 특징 영역은 해당 블록에서 가장 높은 유사도를 가지므로, 이러한 특징영역들의 비교에 의

해 최종적으로 눈 영역을 추출하게 된다. ART2를 통해 후보영역들을 병합하는 과정은 다음과 같다.

[Step 1] 후보 영역들의 특징점을 구하여 입력패턴으로 설정한다.

[Step 2] 새로운 후보 영역의 특징점 X 가 주어진 최소 거리(minimum distance)의 뉴런 j^* 를 승자로 선택하여 표시한다

[Step 3] 경계 임계값 검사(Vigilance test)
승자 뉴런 j^* 가 검사를 통과할 필요충분조건은 다음과 같다.

$$\|X - W_{j^*}\| < \rho$$

- ρ : Cluster 반경

[Step 4] 승자 뉴런 j^* 가 경계 임계값 검사에서 실패하면, 동일한 후보영역부가 아닌 경우이므로 새로운 unit k 가 다음 가중치를 가지고 생성된다.

$$W_k = X$$

[Step 5] 승자 뉴런 j^* 가 경계 임계값 검사에서 통과하면, 즉 가중치 W 와 입력 패턴 X 와의 거리가 경계 임계값 ρ 보다 작으면 동일한 눈 후보 영역부로 인지하고, 승자 뉴런 j^* 의 가중치를 조정한다.

$$W_{j^*}^{(new)} = \frac{X + W_{j^*}^{(old)} \|cluster_{j^*}^{(old)}\|}{\|cluster_{j^*}^{(old)}\| + 1}$$

- $\|cluster_i\|$: $cluster_i$ 에서의 member의 수

위의 ART2 알고리즘에서 Vigilance test를 위한 ρ 값은 20으로 설정하였고 배경의 복잡성에 따라 입력 패턴개수의 차이가 있기는 하였지만 대략 100-200개의 입력 패턴이 생성되었다. 모든 눈 후보 영역 특징점들의 입력이 완료되었을 때 생성되는 클러스터들이 최종 눈 후보 영역이 되며 평균 5-10개가 생성되었다.



[그림 4] 병합된 최종 눈 후보 영역

3.4 눈영역의 판정

신경회로망을 통해 병합된 최종 눈 후보 영역은 기하학적 고찰에 따른 사전지식을 바탕으로 후보영역들간의 위치관계를 고려하여 눈 영역을 판정하게 된다. 최종 눈 후보 영역들은 상에서 하, 좌에서 우 순서로 즉 생성된 클러스터의 인덱스 순서대로 검증을 하게 된다.

먼저, 눈아래의 중심부에 있는 영역들은 대부분 흰색으로 구성되어 있다는 사전 지식을 바탕으로 후보영역들을 검증한다. 이 과정에서 배경영역에서 일부 눈과 비슷한 모양으로 검증된 후보영역이 제거될 수 있다. 또한 눈 영역 추출에서 오검출의 원인이 되었던 눈썹 부분도 이 과정에서 제거된다.

첫 번째 과정을 거치면 눈을 제외한 거의 모든 후보 영역들이 제거되지만, 미세거른 후보영역들에 대해서는 눈의 기하학적 위치정보를 이용하여 눈영역을 검출한다. 두 눈 사이의 간격은 후보영역 폭의 약 2배에서 3배 사이이다. 따라서 현재 검증되는 후보영역은 일정 영역내에 다른 후

보영역이 있어야만 눈 영역으로 인정할 수 있다. 이러한 사전지식을 바탕으로 검증과정을 거치면 고립된 후보영역이 제거될 수 있다. 위에서 언급된 두가지 과정을 거쳤을 때 기하학적 사전지식만으로 대부분 두 개의 눈 영역을 검출할 수 있었다.

그러나, 배경이 복잡하거나 얼굴이 기울어진 영상일 경우 눈의 위치관계만으로는 불충분할 경우가 발생할 수 있다. 즉 위 과정을 거쳤을 때 3개 이상의 후보윈도우가 검출되었을 때는 후보영역들간 유사도 검출 기법으로 최종 눈 후보를 가려낼 수 있다. 먼저 유사도가 가장 높은 윈도우를 눈 영역으로 설정하고, 설정된 눈 영역과 위치정보를 만족하는 후보영역을 두 번째 눈 윈도우로 설정한다. 만약 위치정보를 만족하는 후보영역이 존재하지 않을 경우 존재하는 모든 후보 영역과 가장 근접한 위치관계를 가지고, 동시에 유사도가 높은 후보영역을 최종 두 번째 눈 영역으로 설정한다.



(a) 위치정보만으로 검출된 경우



(b) 위치정보와 유사도 비교로 검출된 경우

[그림 5] 검출된 눈 영역

4. 실시간 눈영역 추적

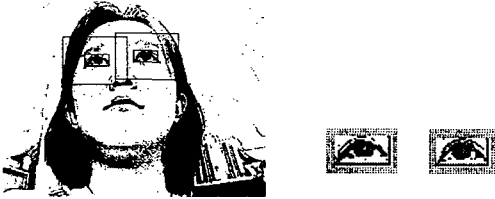
앞에서 언급한 눈 검출 알고리즘은 원영상 전체에 대하여 이미 설정되어진 눈 형판을 이용하여 적용한 것이다. 실시간 처리를 위해서는 최소 초당 15프레임의 영상을 처리해야 된다는 사실을 감안한다면 위의 방법을 사용하여 모든 프레임에 적용한다는 것은 불가능한 일이다. 또한 눈 형판도 고정되어 있으므로 눈의 깜박임과 같은 다양한 변화에 따라 그 정확성도 떨어진다.

따라서, 본 논문에서는 연속적인 얼굴영상에서 첫 프레임만을 얼굴 영상 전체에 대하여, 고정된 형판으로 눈 검출 알고리즘을 적용한다. 만약 첫 프레임에서 눈의 깜박임 등과 같은 변화에 의해 눈 영역 검출 과정이 실패했다면, 다음 프레임에서 같은 과정을 반복한다. 첫 번째 눈 영역이 검출되고 나면 이하의 프레임에서는 다음과 같은 단계로 눈 영역을 검출한다.

- [Step 1] 검출된 눈 영역을 다음 프레임에 위한 형판으로 설정한다.
- [Step 2] 알고리즘 적용 영역을 검출된 눈 영역의 일정 범위내(상하좌우 20pixel)로 한정한다.
- [Step 3] 다음 얼굴 영상 프레임에 입력받는다.

[Step 4] 범위내에서 새롭게 설정된 형판으로 후보영역 추출과 병합, 판정 알고리즘을 적용한다.
 [Step 5] 눈영역 검출 후 [Step 1]으로 돌아간다.

위와 같은 과정을 거치면 알고리즘 적용 범위가 축소되어 처리속도가 매우 향상된다. 또한 이전 프레임에서 검출된 눈 영역 중심으로 다음 눈 영역을 검출하고, 형판 또한 가변적으로 변화하므로 깜박임등과 같은 눈의 변화에도 민감하게 대처할 수 있어 효과적인 눈 영역 추출을 가능하게 한다.



(a) 실시간 알고리즘 적용 영역 (b) 검출된 좌우 눈 영역
 [그림 6] 실시간 알고리즘 적용 영역 및 검출 결과

5. 실험환경 및 결과

실험에서는 Color Sun Video Capture Camera와 시그마 TVII 내장형 캡처 보드를 사용하여 실시간으로 얼굴영상을 캡처하였다. 캡처된 얼굴영상은 320×240 8bit Gray Scale로 구성되는데 카메라로부터 입력되는 컬러영상을 YUV모드로 변환하여 생성하였다. 눈 형판은 30×15의 이진형판으로 설정하였다. 얼굴영상은 안경 착용자와 미착용자 모두를 대상으로 하였고, 조명밝기는 오전, 오후, 저녁등 다양한 밝기에서 실제 촬영하여 실험하였다. 캡처 프레임 속도는 눈 깜박임을 감지할 수 있는 초당 15프레임으로 설정하여 시스템을 구성하였다. 본 알고리즘은 Pentium 450MHz, Memory 128Mbyte, Window98 환경에서 Visual C++ 6.0을 사용하여 구현되었다.

첫 프레임에서 고정 형판을 이용하여 전체영상에서 눈 영역의 검출 알고리즘의 성능을 실험하기 위해, 총 200회 시스템 동작을 하였고, 그 검출 결과는 [표 1]과 같다.

전체동작횟수	검출 성공	검출 실패	성공률	검출시간
200 회	192 회	8 회	96 %	0.17 (sec)

[표 1] 첫 프레임에서의 눈영역 검출 결과

첫 프레임에서의 눈 영역이 검출되고 난 후, 이하 프레임에서는 변형된 형판을 이용하여 설정된 눈 영역 존재 범위내에서만 눈 검출 알고리즘을 적용한다. 그 성능을 실험하기 위해 15 frame/sec의 캡처 속도로 총 5분 동안 4500 프레임의 얼굴영상을 처리한 결과는 [표 2]와 같다.

총 영상	검출 성공	추출 실패	성공률	검출시간
4500	4387	113	97.48 %	1sec/15frm (sec)

[표 2] 두 번째 프레임 이하에서의 눈영역 검출 결과

위의 표에서 보는 바와 같이 본 논문에서 제시하는 알고리즘을 사용하여 눈 영역을 검출한 결과 정확한 눈 영역을 검출할 수 있었으며, 검출시간 또한 적어 실시간 검출이 가능하였다. 위의 표에서 추출에 실패한 경우는 얼굴영상이 30° 이상으로 심하게 기울어진 경우 눈 영역 검출에 실패하였으며, 일부 배경에 눈과 매우 유사한 형태가 감지되었을 경우 첫 프레임에서의 눈 영역 검출에 실패하

였다. 얼굴이 심하게 기울어진 경우 두 눈의 특징점을 설정하여 중점을 계산하는 방법 또는 분할 방법으로 기울기를 조정하면 기울기 문제는 해결할 수 있을 것으로 보인다.

6. 결론

본 논문에서는 명암의 변화에도 비교적 적응력이 강한 이진화 기법을 사용하여 원 영상을 이진화하는 방법을 제시하였다. 또한 가변 템플릿 방법을 사용한 후보 영역 추출방법과 ART2 신경회로망을 이용한 병합방법을 제안하였으며, 병합된 후보 영역들을 얼굴 요소의 기하학적 사전 지식을 기반으로 검증하는 방법을 설명하였다. 이러한 방법은 실험결과에서 보느바와 같이 다양한 배경을 가진 얼굴영상에서 시간에 따라 모양변화가 급변하는 눈 영역에 대한 실시간 추출을 가능하게 하였다. 이상의 연구 결과는 환자와의 의사소통 및 감시장치 시스템에 적용이 가능하며, ITS(Intelligent Transportation System)분야에도 적용되어 교통사고를 방지를 위한 눈의 출력감지 등의 응용 시스템에 이용될 수 있다.

향후 연구과제로는 앞에서 언급되었던 눈의 기울기 처리 문제와, 가장 큰 오류의 원인이 되는 국부 조명의 변화에 능동적으로 대처하는 방법을 개발하는 것이다. 또한 환자와의 의사소통 및 줄음감지를 위해 검출된 눈 영역에서의 눈의 개폐 유무를 판단할 수 있는 방법을 연구해야 한다.

참 고 문 헌

- [1] Henry A. Rowley, Shumeet Baluja, Takeo Kanade, "Rotation Invariant Neural Network-Based Face Detection", 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 23-25, Santa Barbara, California, pp.38-44.
- [2] 김성환, 한영환, 박승환, 장영진, 홍승홍, "환자 감시장치를 위한 눈의 개폐상태 인식에 관한 연구", 대한전자공학회, B-1995, 11v.32-B, n.11, 1995, pp.1455-1463.
- [3] 윤호섭, 왕 민, 민병우, "눈 영역 추출에 의한 얼굴 기울기 교정", 대한전자공학회, v.33-B, n.12, 1996, pp.71-83.
- [4] 풍의섭, 김병화, 안현식, 김도현, "부계중심을 이용한 자동얼굴인식 시스템의 구현", 대한전자공학회, v.33-B, n.8, pp.114-123.
- [5] Rafael C. Gonzalez, Richard E. Woods "Digital Image Processing", Addison Wesley, pp.443-455, 1992.
- [6] Milan Sonka, Vaclav Hlavac and Roger Boyle, "Image Processing Analysis and Machine Vision", University Press, Cambridge, pp.113-121, 1993.
- [7] Ramesh Jain, Rangachar Kasturi, Brian G. Schunck, "Machine Vision", McGraw-Hill, pp.76-86, 1995.
- [8] Michael Chester, "Neural Networks", Prentice Hall, pp.71-81.