

## 사용자 인터페이스 기능의 개인용 3D 방송단말을 위한 얼굴 및 눈인식 알고리즘의 고속 검출 방법

\*송혁    \*\*곽수영    \*\*\*배진우    \*\*\*\*최병호

\* \*\*\*\*전자부품연구원, \*\*한밭대학교, \*\*\*R&D 센터

\*hsong@keti.re.kr, \*\*\*\*bhchoi@keti.re.kr, \*\*sykwak@hanbat.ac.kr, \*\*\*bjw8751@kipa.org

### High Speed Face and Eye Detection Algorithm for Personal 3D Broadcasting Devices with User Interactive Function

\*Song, Hyok    \*\*Kwak, , Soo-Yeong    \*\*\*Bae, Jin-Woo,    \*\*\*\*Choi, Byeong-Ho

\* \*\*\*\*Korea Electronics Technology Institute, \*\*Hanbat University, \*\*\*R&D Patent center

#### 요약

3D 디스플레이 장치는 디스플레이 장치의 시점 수에 따라 다시점 디스플레이 장치 및 스테레오 디스플레이 장치로 구분할 수 있다. 스테레오 디스플레이 장치는 좌우 두 개의 영상을 디스플레이 방식에 따라 영상포맷을 변경하여 디스플레이에 보여지고 다시점 디스플레이는 두 개 이상의 영상 데이터를 포함하는 것으로 구분할 수 있다. 영상의 화질문제로 인하여 최근 3DTV는 스테레오 방식을 사용하고 있으며 대기업에서 양산하고 있는 기술이다. 스테레오 방식에서는 개인용 단말의 경우 사용자의 위치 및 눈의 위치에 따라 깊이감을 달리 보여주는 장치가 개발되고 있어 이에 대한 연구를 진행하였다. 이와같은 기술의 흐름에 맞춰 3D 디스플레이 장치를 위한 눈 추적 알고리즘을 구현하였다. 본 논문에는 카메라에서 입력되는 영상에서 얼굴영역을 Haar training 방법을 통하여 추출하였으며 PCA 방법을 통하여 눈 영역을 추출하였다. 알고리즘의 복잡도로 인하여 연산량을 줄이기 위하여 Downsampling 방법을 활용하여 연산량을 약 26% 감소시켰다. 본 연구 결과를 통하여 얼굴 및 눈 추적 알고리즘 구현시 최적화를 위하여 필요한 프로세스를 예측할 수 있다.

#### 1. 서론

아바타로 시작된 3D 관련 기술은 그 범위가 폭발적으로 늘어나고 있다. 3D 콘텐츠 제작기술, 전송기술, 디스플레이 기술 뿐 아니라 관련 기기 분야로까지 시장이 확대되고 있다. 이제 3D 기술은 디스플레이 관련 기기에서 기본 사양으로 채택되어가고 있다. 그러나 실제 3D 디스플레이 장치는 스테레오 방식의 디스플레이가 주도하고 있다. 양산되고 있는 디스플레이 장치의 경우 스테레오 방식으로 서터글라스 방식과 편광방식의 3DTV가 시판되고 있다. 그러나 디스플레이 장치에서 보여지는 영상을 사용자가 관찰시점을 좌우로 이동하거나 앞뒤로 이동할 경우 3D에서 발생하는 깊이감 및 실감도가 줄어들거나 혹은 너무 큰 깊이감으로 인하여 피로를 발생할 수 있다. 특히 렌티큘라방식이나 패럴랙스배리어 방식의 스테레오 디스플레이 장치는 관찰자의 위치에 따라 3D 실감정도를 느끼는데 큰 차이를 보인다. Autostereoscopic 디스플레이는 사용자의 위치를 파악하고 그에 따라 디스플레이되는 데이터를 변화함으로 위치에 상관없이 가장 적절한 3D 실감영상을 제시할 수 있다. 기존 Autostereoscopic 디스플레이를 구현하기 위해서는 적외선센서나 초음파센서와 같은 센서장치를 부착하는 연구도 있으나 오히려 사용자의 움직임을 제한하는 문제가 발생한다. 따라서 영상 정보를 활용한 Interactive 3D 디스플레이장치가 개

발되었으며 주로 크로스토크를 없애기 위해서 연구되었다[1]. 최근에는 사용자의 얼굴 및 눈의 위치를 추적하여 위치 뿐 아니라 로테이션을 감안한 영상변환도 고려하였다.

#### 2. 본론

얼굴 및 눈 추출 알고리즘은 그림 1과 같이 전처리부, 얼굴인식부, 얼굴영역 추출부, 얼굴영역 전처리부, 눈검출부로 구분된다. 영상을 활용한 인식 시스템은 조명의 변화에 매우 민감하므로 얼굴영역의 전처리는 매우 중요하다.

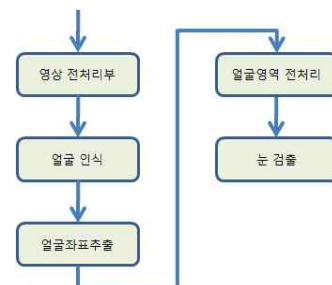


그림 1. 눈 검출 프로세스

디스플레이 장치에 연동하여 사용자의 위치정보를 추출하여 반응하기 위해서는 처리 속도가 매우 중요하다. 최초 눈인식 알고리즘의 구현 속도는 5프레임/Sec의 인식 알고리즘을 OpenCV를 활용한 Visual studio를 활용하여 구현하였으며 이를 활용 가능한 인식속도로 구현하기 위하여 20프레임/Sec를 목표로 하였다. 이를 위하여 전체 프로세스 중에서 각각의 모듈에서 사용하는 프로세싱 시간을 연산하였다.

표 1. 검출 시간

프로세스	얼굴검출	왼눈검출	오른눈검출	후처리 및 디스플레이
연산시간 (msec)	0.56	16.02	30.10	55.30

표 1과 같이 검출 시간을 추출할 수 있었으며 각 시간은 최초 프로세스 시간에서부터의 시간을 의미하며 좌우 눈을 검출하는데 각각 약 15 msec 정도가 필요하였다. 좌우 눈의 검출 소요시간이 상이한 것은 소프트웨어 코드상에서 눈 검출을 위하여 선 작업 및 전처리과정을 포함하기 때문이다. 따라서 영상 디스플레이 과정을 제외한 전체 프로세스에서 눈검출 시간이 55%정도 차지함을 알 수 있었다. 본 논문에서는 눈 검출을 위해 얼굴영역 내에서 눈이 존재하는 후보영역을 제한하여 연산량을 줄였음에도 가장 많은 시간이 소요되었으므로 검색시간을 줄이기 위하여 Downsampling 방법을 적용하였다. 초기 입력 영상은 320x240 크기의 영상이었으며 얼굴영역의 크기는 너비 100 픽셀 전후의 크기이며 눈 크기는 너비 18픽셀 전후의 값을 보였다. 좌우 후보 영역에서 눈 영역을 찾기 위한 위치정보는 좌우25픽셀 전후 및 상하 17픽셀 전후의 영역으로 평균 425번의 연산을 통하여 하나의 눈 영역을 추출하였다. 이를 눈 후보 영역을 Downsampling 후에 Distance가 가장 적은 영역을 추출한 후 주변 영역을 검색하였을 경우 연산량은 아래와 같다.

표 2. Downsampling 후 검출시간 비교

	얼굴 검출	왼눈 검출	오른눈검출
1BLK	1.11	17.40	31.17
2BLK	2.32	15.74	21.31
3BLK	2.07	15.61	23.04
4BLK	3.31	18.56	36.54
5BLK	4.18	33.66	51.62

표 3. 눈 검출 알고리즘에서의 좌우 Distance 값의 비교

	distanceL	distanceR
1BLK	416.16	349.65
2BLK	414.05	351.01
3BLK	413.1	347.74
4BLK	413.27	350.40
5BLK	412.31	349.47

표 2에서 보는 바와 같이 Downsampling 블록의 사이즈에 따라 연산량이 감소하는 것을 볼 수 있었다. 표 3은 눈 검출시 Distance 값을 나타낸다. 표에서 보는 바와 같이 Downsampling 비율에 관계없이 유사한 Distance 값을 나타냄을 볼 수 있다. 그러나 4블록 Downsampling 이상 진행하게되면 1차 검색 영역 주변에서 다시 원영역 블록 주변에서 검색하여야 하는 블록 수가 많아져 연산속도가 떨어짐을 볼 수 있었다. 이와 같은 결과는 Downsampling시 발생하는 오류를 줄이기 위한 8-neighbour 검색 알고리즘에 의한 것으로 그림 2와 같이 Downsampling 검색 후 8-neighbour 블록을 검색한다.

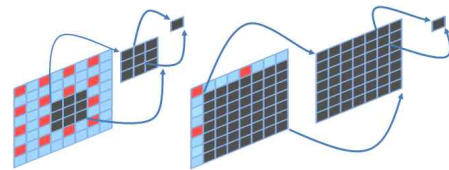


그림 2. 8-neighbour 검색 방법

본 논문은 WindowsXP OS 및 i7 CPU가 장착된 시스템에서 얻어진 연구결과이나 절대적 계산값은 달라질 수 있으나 상대적 소요시간은 다른 시스템에서도 같을 것으로 예상된다.

### 3. 결론

본 논문은 개인용 3D 단말기에서 사용자가 단수일 경우 적절한 깊이감 및 최적의 3D 실감 정보를 전달하기 위하여 사용자의 얼굴 및 눈 위치를 추적하여 3D 정보를 변형하기 위한 위치추출 알고리즘으로 고속 위치 추적 알고리즘을 구현한 것이다. 검출 영역의 제한만으로 총 26% 연산량을 감소시켰다. 연산량 감소에도 불구하고 각 Downsampling 방법에 Distance 값의 변화가 거의 없었다. 그러나 연산 시간에 차이가 있으므로 2BLK Downsampling 방법 및 8-neighbour 방법을 적용한 알고리즘이 가장 효율이 좋을 수 있었다.

본 연구결과에서 얼굴 및 눈 검출 알고리즘에서 발생하는 연산시간의 비중이 매우 컸으나 영상획득 및 디스플레이부의 연산시간 역시 매우 큰 부분을 차지하는 것으로 나타나 이 부분에 대한 소프트웨어적인 최적화가 뒤따라야 한다는 것을 알 수 있었다.

본 연구 결과는 각 Downsampling 검색시 동일한 282개의 프레임을 적용하여 평균값을 구한 것으로 WindowsXP에서 동작하였으므로 각각의 시간에서의 오차가 발생하는 경우가 있었다.

[1] G. J. Woodgate, D. Ezra, J. Harrold, N. S. Holliman, G. R. Jones, R. R. Moseley, "Autostereoscopic 3D display systems with observer tracking", Signal Processing-image Communication - SIGNAL PROCESS-IMAGE COMMUN , vol. 14, no. 1-2, pp. 131-145, 1998