

## 2채널 3차원 입체음향 구현을 위한 머리 위치 검출

이상준<sup>o</sup>, 박장식, 손경식

부산대학교 전자공학과, 동의공업대학 영상정보과,

### Head Tracking for Implementing 3-Dimensional Stereo Sound

Sang Jun Lee, Jang Sik Park and Kyung Sik Son

Dept. of Electronics Eng. Pusan National University

L2516@hanmail.net

Dept. of Visual Technology, Dongeui Institute of Technology

jsipark@dit.ac.kr

#### 요약

입체음향을 구현하는 방법은 5.1 채널을 이용하는 방법과 사람이 두 귀를 이용하여 정위감을 느끼는 원리와 같은 2 채널 입체음향을 구현 방법이 있다. 2 채널 입체음향을 재생하는 것은 5.1 채널에 비하여 설치하는 측면에서 효율적이지만 cross-talk 를 제거하는 것이 어려운 문제이다. Cross-talk를 제거하기 위해서는 머리의 위치를 정확하게 추정하는 것이 필수적이다. 본 논문에서는 2 채널로 3 차원 입체음향을 재생하기 위하여 머리의 위치를 추정하기 위한 알고리즘을 제안한다. 제안하는 알고리즘은 전처리를 통하여 머리의 범위를 검출하고 수학적형태학을 이용하여 눈의 위치를 검출하여 머리가 향하고 있는 방향을 검출한다. 본 논문에서 제안하는 알고리즘이 효율적으로 머리의 위치를 검출할 수 있음을 시뮬레이션을 통하여 보여준다.

#### 1. 서론

신호처리 및 DSP 기술의 발전으로 멀티미디어 처리 기술이 현저히 발전하였다. 멀티미디어 기술에 있어 비디오 뿐만 아니라 현장감을 제공하기 위해서는 입체음향을 재생할 수 있는 오디오 기술들이 요구되고 있다.

3차원 입체음향을 재생하는 방법은 크게 두 가지가 있다. 기본적으로 5 개의 스피커와 저음을 재생하기 위한 5.1 채널이 일반적으로 널리 사용되고 있는 방법이다. 그러나 5.1 채널로 입체음향을 재생하기 위해서는 6 개의 스피커를 설치하여야 하며 헤드셋을 착용하는 경우에는 이를 실현하는 것이 어렵다. 3차

원 입체음향을 재생하는 방법으로 사람의 청각 특성과 동일한 방법을 사용하는 것이 2 채널 HRTF(head related transfer function)을 이용하는 방법이 있다[2]. HRTF 를 이용하여 스피커로 입체음향을 구현하는데 있어 가장 큰 장애 요소인 Cross-talk를 제거하여야 한다. Cross-talk를 제거하기 위해서는 머리의 위치를 검출하여 역모델을 적용하는 것이다 일반적이다[1]. 본 논문에서는 카메라를 이용하여 머리의 위치를 검출하기 위한 방법을 제안하고자 한다. 머리의 위치를 검출하는 방법은 다양하게 제안되었지만 본 논문에서는 실시간으로 머리의 위치 검출을 위하여 밝기 정보를 이용하여 얼굴의 영역을 검출하고 수학적형태학을 이용하여 눈의 위치를 검출하여

머리가 향하고 있는 방향을 검출하는 방법을 제안한다[3]. 본 논문에서는 컴퓨터 시뮬레이션을 이용하여 제안하는 방법이 유용함으로 보인다.

## 2. 2채널 입체음향에서의 Cross-talk

그림 1과 같이 2 채널로 입체음향을 라우드 스피커로 재생하고자 한다면 헤드셋을 착용하는 것과 달리 스피커의 음원과 청취자 사이에 cross-talk 가 발생하게 된다[1].

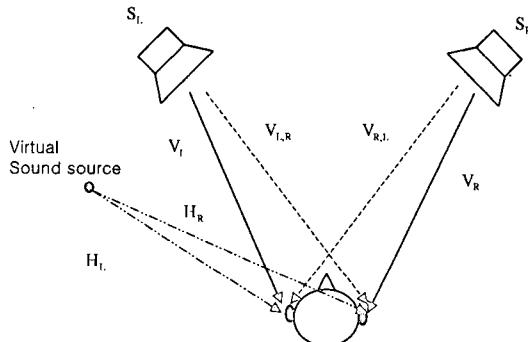


그림 1. 2 채널 입체음향 재생에 있어 cross-talk

## 3. 머리 위치 검출 알고리즘

본 논문에서 제안하는 머리의 위치 검출 알고리즘은 그림 2와 같다. 카메라로부터 입력되어진 영상밝기 이용하여 얼굴의 영역을 검출하고 수학적형태학을 이용하여 눈의 위치를 검출한다. 얼굴의 영역과 눈의 위치를 분석하여 최종적으로 머리가 향하고 있는 방향을 결정한다.

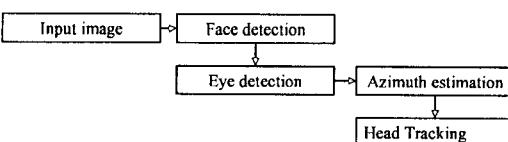


그림 2. 제안하는 머리 위치 검출 알고리즘

### 3.1 얼굴 영역 검출 알고리즘

그림 3은 입력영상으로부터의 얼굴 영역을 검출하기 위한 알고리즘을 나타내고 있다. 얼굴 영역을 검출하기 위해선 주어진 영상의 밝기 정보를 이용하여 얼굴의 가로, 세로의 임계값을 구하여 얼굴의 영역을 검출한다[3].

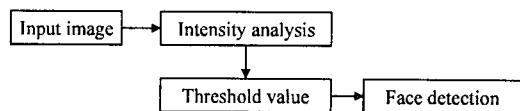


그림 3. 얼굴 영역 검출 알고리즘

그림 4는 가로 및 세로 방향으로 평균 밝기값을 구하여 얼굴 영역을 구하는 과정을 보여주고 있다. 가로 영역은 얼굴 정면의 밝기에 대하여 좌우의 밝기가 25 % 정도 된다고 가정하고 정면의 밝기에 대하여 25 %의 밝기의 지점을 얼굴 좌우의 끝점으로 한다.

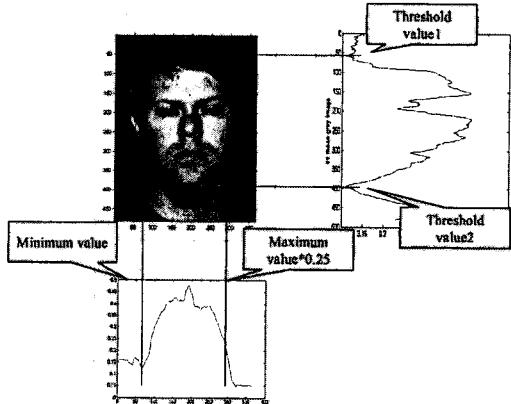


그림 4. 얼굴위치 검출 방법

그림 4와 같이 얼굴의 세로 영역은 좌우에 비하여 평균 밝기값의 분포가 다양하게 나타난다. 이것은 얼굴에서 눈썹, 눈 그리고 입이 있는 위치에서 각각 어두운 밝기값을 나타내기 때문이다. 세로로 10 픽셀 단위에서 최대값을 구하고 그 중에서 3 개의 최대값을 순서대로 구한다. 그리고 위에서부터 첫번째 최대값사이의 최소값을 세로 영역 윗부분으로 설정하고 3 번째 최대값에서부터 아랫 부분까지의 최소

값을 얼굴의 아랫부분으로 설정한다.



그림 5. 얼굴 영역 검출 결과

그림 5는 평균 밝기값을 이용하여 얼굴 영역을 검출한 결과이다.

### 3.2 눈 위치 검출 알고리즘

그림 6은 눈의 위치 검출하는 순서를 나열한 것이다. 입력된 계조영상에 대하여 히스토그램을 이용하여 이진영상으로 변환하고 수학적형태학을 이용하여 눈의 위치를 검출한다[4].

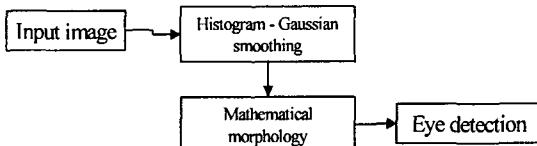


그림 6. 눈의 위치 검출 알고리즘의 구성도

수학적형태학은 영상분석, 영상분리 등의 다양한 분야에 활용되고 있는 영상처리 방법으로써 영상의 형태를 기반으로 dilation, erosion, complementary 연산을 기반으로 다양한 영상처리를 할 수 있다.

그림 7은 히스토그램을 이용하여 이진영상으로 바꾸는 과정을 나타내고 있다. 그림 8은 이진영상으로부터 수학적형태학을 기반으로 눈의 위치를 검출하는 과정을 나타낸다.

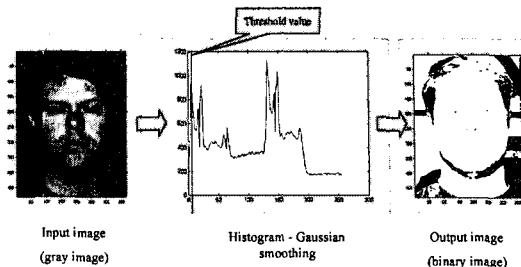


그림 7. 계조영상을 이진영상으로 변환

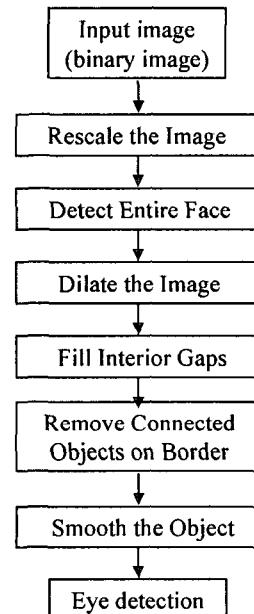


그림 8. 수학적 형태학을 이용한 눈의 위치 검출 알고리즘



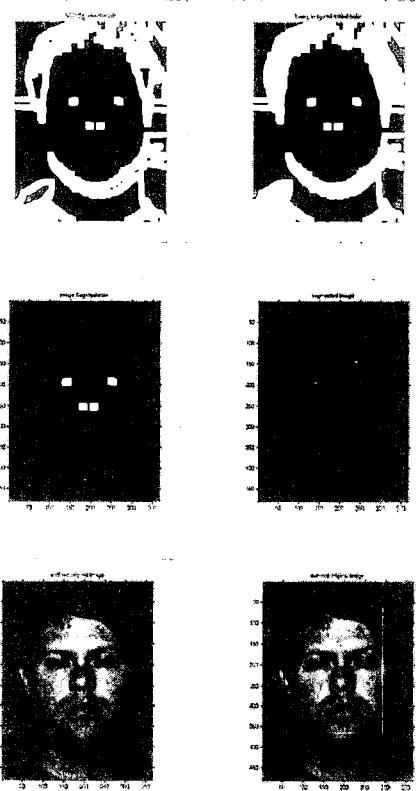


그림 9. 눈의 위치 검출하기 위하여 단계별로 처리된 결과

### 3.3 각도 추정

앞에서 구한 얼굴 영역과 눈의 위치를 이용하여 얼굴의 정면이 어느 방향으로 향하고 있는지 검출할 수 있다. 얼굴이 정면을 향하고 있을 때 얼굴영역의 왼쪽 끝점에서 왼쪽 눈의 중심 위치까지 74 픽셀이었고 얼굴이  $12^\circ$  오른쪽으로 향하고 있을 때는 90 픽셀이었다. 얼굴이 전체적으로  $24^\circ$  오른쪽으로 향하고 있을 때는 120 픽셀 거리로 계산되었다.

## 4. 결론

본 논문에서는 밝기 정보를 이용하여 얼굴 영역을 검출하고 수학적 형태학을 이용하여 눈의 위치를 검출하여 최종적으로 머리가 향하고 있는 방향을 검출

하는 방법을 제안하였다. 제안하는 방법은 2채널로 3차원 입체 음향을 구현하는데 있어 cross-talk 를 제거하는데 효과적일 것으로 판단된다. 제안하는 방법에서는 일반 가정의 거실을 중심으로 오디오 및 비디오 시스템이 배치되는 환경을 고려하여 조명이 일정하게 정면에 있는 것으로 가정하였지만 보다 넓은 공간에서는 위의 가정이 적용되지 않을 것이다. 따라서 불균일한 조명환경에 대하여 머리의 방향을 검출하는 방법이 필요하며 다양한 얼굴 영상으로 얼굴이 정면을 기준으로 향하고 있는 방향과 픽셀간의 관계를 결정하는 것이 필요하다.

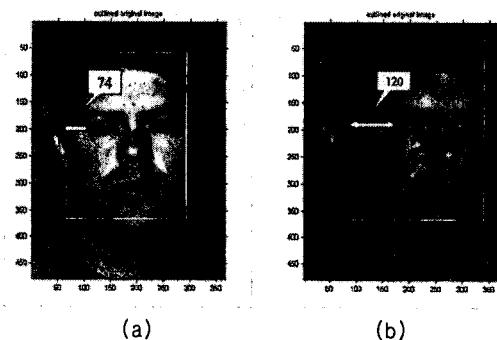


그림 10. 머리 영역과 눈의 위치를 검출한 결과

- (a) 머리가 정면을 향하고 있는 영상
- (b) 머리가  $24^\circ$  오른쪽을 향하고 있는 영상

### [참고문헌]

- [1] H. S. Kim, P. M. Kim and H. B. Kim, "Cross-talk Cancellation Algorithm for 3D Sound Reproduction," ETRI Journal, pp. 11-19, VOL. 22, NO. 2, June, 2000
- [2] Chris Kypriakakis, "Fundamental and Technological Limitations of Immersive Audio Systems," Proceedings of IEEE, pp. 941-952 VOL. 86, NO. 5, May, 1998
- [3] Mohamed Rizon and Tsuyoshi Kawaguchi, "Automatic eye detection using intensity and edge information," IEEE, pp. 415-420, VOL II, 2000
- [4] Liang Tao and H. K. Kwan, "Automatic localization of human eyes in complex background," IEEE, pp. 669-672, VOL V, 2002