

음원인식 및 지연시간을 이용한 카메라의 방향제어 시스템 설계

이희태*, 김영섭**

*동아대학교 전자공학과

** (주)시그널웍스

e-mail : lht3318@naver.com

Design of direction control system for camera, Using sound source recognition and delay time.

Hui-Tae Lee*, Young-Sub Kim**

*Dept of Electronic Engineering, Dong-A University

** (Corp)SignalWorks

요 약

본 연구는 이상음원(비명, 유리창 파손음, 경적소리 등) 발생 시, 2개의 마이크로폰에 입력되는 사운드에 대하여 음원 방향추적 장치와 연결된 카메라에 음원의 방향 정보를 전송함으로써, 카메라의 View Point를 음원 발생방향으로 이동시켜 사고현장을 더욱 신속하게 대처할 수 있는 시스템에 대한 연구이다. 일반적인 음성을 이용한 감시카메라는 단순히 소리 발생 여부만 감지하지만, 본 시스템은 이상음원 발생 지점으로 카메라의 방향 제어를 가능하게 한다. 이상음원의 검출은 기존에 수집한 DB를 기반으로 비교, 분석 과정을 통하여 이상음원을 분류한다. 음원 발생 방향은 음원 발생 시, 마이크로폰에 도달하는 음원의 시간차에 따른 음파의 위상차를 계산하여 음원 발생 방향을 판단하게 된다.

1. 서론

최근 음성인식 기술은 현대 사회의 많은 분야에 적용되어 사용되고 있다. 특히, 보안 시스템에도 이러한 음성인식 기술을 적용할 수 있으며, 최근 실제로도 많은 연구 및 개발이 이루어지고 있다. 이에 따라, 보안 시스템 중 하나인 감시카메라에도 음성인식 기술의 적용이 늘어나는 추세이다.

하지만 대부분 개발되어 있는 음성을 이용한 보안 장비들은 단순히 소리 발생 여부만 감지하며, 음원 발생 지점으로 카메라의 방향을 제어할 수가 없다. 이는 비정상 상황 시 발생하는 소리에 대해 발생 여부만 알 수 있고, 그 소리가 무엇 때문에 발생했는지에 대해서는 규정하기 힘들다. 따라서 본 시스템 설계로 인하여 음원 발생 지점으로 카메라의 방향 제어를 가능하게 함을 목적으로 한다.

본 시스템은 오디오에 수신되는 음향 신호가 비정상 음원으로 판단 될 경우, 음원에 대한 방향 추적을 진행한다. 또한, 2개의 마이크로폰으로 수신되는 신호 간의 도착 시간 차이를 상관도를 이용해 계산하여, 음원 발생 방향을 추적하는 시스템이다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 본 시스템 설계 과정 중 이상음원 비교, 분석에 관하여 서술하며, 3장에서는 음향신호를 이용한 방향 검출 기술에 관하여 서술한다. 4장에서는 본 시스템의 구성에 관하여 설명하며, 마지막으로 결론을 기술하며 끝을 맺는다.

2. 이상음원 비교, 분석

먼저, 마이크로폰에 이상음원(비명, 유리창 파손음, 경적소리 등)이 감지 될 경우, 음원 전개 및 변화의 특성, 음량의 변화 패턴, 특정 주파수의 성분비 분석 등을 이용하여 비명, 유리창 깨짐, 차량 충돌 등으로 분류한다. 그 후 이상음원 검출기 내에서 이상음원을 분석할 수 있는 기준 음원 DB 비교 기술을 통해 음원 분석을 진행하게 된다.

이러한 사운드 수집 및 처리는 그림1의 Raspberry Pi를 이용하여, Open source hardware를 이용한 사운드 수집 및 처리 과정을 거친다.



그림 1. Raspberry Pi

3. 음향신호를 이용한 방향 추적

3.1 신호간의 도착 지연시간 계산

음원의 지연 시간을 구하는 방법에는 TDOA(Time Delay of Arrival), 그리고 GCC-PHAT(Generalized Cross Correlation Phase Transform)등이 있다. TDOA는 상대적

으로 계산량이 적고 비교적 정확성이 높으나, 주변 잡음 및 여러 요소에 따른 영향을 받기 쉽다. 반면 GCC-PHAT의 경우, 계산량은 상대적으로 많으나, 잡음이 존재하는 환경에서 강인하다는 장점이 있다. 따라서 본 시스템에서는 음원의 위치 추적을 위해 GCC-PHAT를 이용하며, 신호 간의 도착 지연시간을 계산하기 위해 식(1)과 (2)와 같이 계산한다.

$$R_{x_1x_2}(n) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W(w)X_1(w)X_2^*(w)e^{jwn} dw$$

$$W(w) = \frac{1}{|X_1(w)X_2^*(w)|} \quad (1)$$

$$\tau = \arg \max R_{x_1x_2}(n) \quad (2)$$

식(1)의 $W(w)$ 는 주파수 가중 함수이고 $X_1(w)X_2^*(w)$ 역수이며, 이 가중 함수를 PHAT라 한다. PHAT는 시간 지연을 추정함에 있어, 수신되는 신호의 주파수에 종속된 가중치 된 함수이다. 즉, 각 주파수의 상대적 중요성을 결정하며 이를 식(2)에 나타내고 있다.

식(1)을 통해 구해진 $R_{x_1x_2}(n)$ 을 이용해 최종적인 지연시간을 식(2)과 같이 구할 수 있다.

3.2 지연 시간을 이용한 음원 방향 추적

이상음원 발생 시, 2개의 마이크로폰에 도달하는 음원의 지연시간을 상기의 방법을 이용하여 계산한 후, 이상음원 발생 방향을 추적한다. 그림2는 음원으로부터 각 마이크로폰에 입력되는 거리에 따른 지연시간을 이용하여 음원 발생 방향을 추적하는 과정을 보인 것이다.

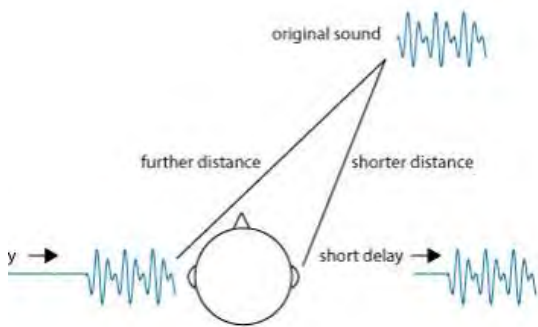


그림 2. 음원 방향추적 과정

4. 시스템 구성도

그림3은 본 시스템의 구성을 간략하게 나타낸 것이다.

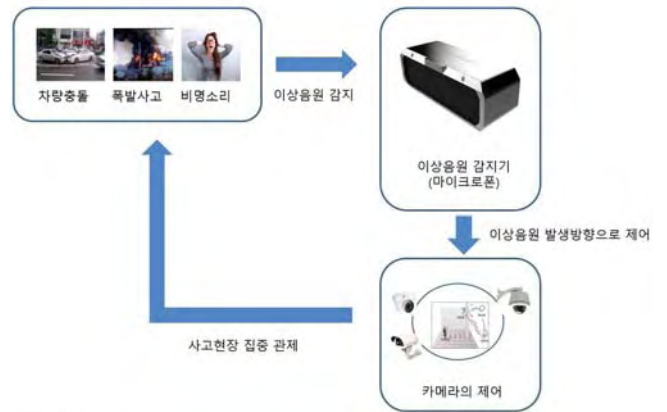


그림3. 시스템의 구성도

- ① 사건이 발생하면 Raspberry Pi를 이용한 음성인식 센서로 이상음원을 감지한다.
- ② 이상음원 검출기 내에서 기준음원 DB 비교기술을 통한 음원 분석 기술로 이상음원임을 판단한다.
- ③ 마이크로폰에 도달하는 음원의 시간차를 계산하여 음원 발생 방향을 판단한다.
- ④ 이상음원의 방향이 감지가 되면, 카메라의 View Point를 이상음원 발생지로 이동한다.

5. 결론

본 연구는 이상음원(비명, 유리창 파손음, 경적소리 등) 발생 시, 마이크로폰에 입력되는 사운드에 대하여 음원 방향추적 장치와 연결된 카메라에 음원의 방향 정보를 전송함으로써, 카메라의 View Point를 음원 발생 방향으로 이동시켜 사고현장을 더욱 신속하게 대처할 수 있는 시스템에 대하여 연구했다. 기존 보안 시스템의 Hardware는 공간과 비용 측면에서 단점이 있었으나, 이를 보완하기 위해 Raspberry Pi를 이용하여 실험을 진행하였다. 하지만 Raspberry Pi는 실제 고가의 서버용 컴퓨터보다는 용량과 처리속도의 한계가 있어, 음원의 세밀한 분석 관점에서는 Raspberry Pi가 뒤떨어지는 문제점도 발생하였다.

본 연구와 같이, 인식을 통해 학습을 하는 시스템의 경우, 표본의 개수에 비례해 인식률이 증가하게 된다. 따라서 본 시스템 또한, 표본의 개수가 더욱 증가하고, Raspberry Pi와 같은 H/W의 성능과 단점이 개선된다면, 정확도는 더욱 증가할 것이며 다양한 분야에 적용될 수 있을 것으로 생각한다.

참고문헌

[1] L. R. Rabiner and R. W. Schafer, "Digital Processing of Speech Signals, Prentice-Hall," 1978.

[2] P. Atrey, N. Maddage, and M. Kankanhalli, "Audio Based Event Detection for Multimedia Surveillance," 2006 IEEE Int. Conf. on Acoustics Speech and Signal Processing Proceedings, pp. 813-816, Toulouse, France, May 2006.

[3] J. R. Doller, J. G. Proakis and J. H. L. Hansen, "Discrete-Time Processing of Speech Signals," Macmillan, 1993.

[4] H. Wang and P. Chu, "Voice source localization for automatic camera pointing system in video conferencing," Acoustics, Speech, and Signal Processing, ICASSP-97, vol. 1, pp. 187-190, 1997.

“본 논문은 2017년 한이음 ICT멘토링 프로젝트의 결과물입니다.”