

# 지능형 감시 시스템을 위한 음원 위치 파악 및 침입자 추적 시스템

## Sound localization and object tracking for intelligent surveillance system

박정현<sup>1</sup>, 정봉규<sup>1</sup>, 이민영<sup>2</sup>, 심귀보<sup>1</sup>

<sup>1</sup> 중앙대학교 전자전기공학과

<sup>2</sup> (주)바이오인식

(E-mail: kbsim@cau.ac.kr)

### 요 약

보안이 중요시되는 공간에서 임의의 사람을 추적하고 인식할 수 있는 시스템의 필요성이 점차 중요시되고 있다. 본 논문에서는 wide-area를 감시하기 위해 음원과 Pan-Tilt 카메라를 이용하여 침입자 추적 시스템을 구현하고자 한다. 먼저 음원 추적 센서를 이용하여 음원의 방위를 검출한 후 Pan-Tilt 카메라를 방위 방향으로 이동시키고 이동된 Pan-Tilt 카메라로부터 기준영상을 추출한 후 일정 기준 시간마다 영상을 다시 획득하여 차 영상과 검출결과를 얻어 침입자를 추적하여 감시 시스템을 구현한다. 본 논문에 의해서 구현된 시스템은 홈 네트워크 보안 시스템과 연동하여 가정이나 공공시설 보안에 대한 인터페이스 역할을 할 수 있다.

**Key Words** : 지능형 감시시스템, 음원추적 센서, Pan-Tilt 카메라, Surveillance

### 1. 서 론

9.11 테러 이후 공공시설, 기업, 가정 등에서 보안에 대한 중요성이 커지고 있다. 이에 따라 효과적인 보안 시스템 구축이 필요하다. 보안 업체는 인체 감지 센서(PIR)나 마크네틱 센서를 이용하여 침입 감지를 인지하고 데이터를 PSTN(Public Switch Telephone Network)망을 통하여 서버에 전송하고 있다. 위와 같은 시스템은 사용자 부주의나 센서의 오동작에 취약할 뿐 아니라 침입자에 대한 정보를 가지고 있지 않기 때문에 보안 시스템으로 취약성을 보이고 있다. 이에 대한 대안으로 제일 효과적인 것은 영상을 이용한 보안 시스템이다.

인터넷 망의 발달로 인하여 영상 전송이 가능해 짐에 따라 시스템 구축이 가능해 졌다. 그러나 고정 카메라를 이용하여 영상을 저장하고 전송하는 것은 유효 영상을 얻기 힘들다. 이에 여러 개의 고정 카메라와 Pan-Tilt 카메라를 조합하여 네트워크를 구성하는 방법이 제

안되었다[1]. 이 방법은 유효 영상을 얻기 위해서는 효과적이지만 여러 대의 카메라를 필요로 할뿐만 아니라 여러 개의 영상을 서버로 전송할 데이터가 많아지기 때문에 보안 시스템 구축에 이용에는 적절치 않다. 또한, 고정 카메라와 PTZ(Pan-Tilt-Zoom)카메라를 이용하여, 고정 카메라로 움직임을 감지하고, PTZ카메라를 이용하여 움직임을 추적함으로써 영상을 획득하는 방법이 제안되었다[2]. 이 방법 또한 2개 이상의 카메라가 필요하며, 실내나 빌딩 내부에 적용하기에는 시스템이 커지는 단점이 있다.

본 논문은 한 개의 PT(Pan-Tilt) 카메라를 이용하여 wide-area를 감시하기 위한 방법으로 음원을 이용하고자 한다. PT카메라 한 개를 이용할 경우에는 현재 위치한 획득 영상에 대해서만 감지를 할 수 있기 때문에 카메라의 Pan-Tilt 기능을 이용하여 움직임에 대하여 검색을 하여 유효 영상을 획득해야 한다. 본 논문은 효과적인 시스템 구축을 위하여 카메라를 이용해 검색하는 방법대신 음원추적 센서를 이용 하였다. 음원의 방향으로 카메라를 이동하여 침입자를 검색한 후 이동 경로에 따라 카메라를 이동하여 침입자의 감지 영상을 저장하는

감사의 글 : 본 연구는 산업자원부 성장동력산업과제의 연구비 지원으로 수행되었습니다. 연구비 지원에 감사드립니다.

시스템을 구현한다. 본 논문의 2장에서는 음원 추적 센서를 이용한 음원의 위치 인식에 대해 기술하고, 3장에서는 PT Camera를 이용한 움직임 추적에 대해 기술하고, 마지막으로 4장에서는 실험 및 그 결과에 대해 기술한다.

## 2. 음원 추적 센서를 이용한 음원의 위치 인식

기존의 음원추적 센서를 설계하는 방법으로 서로 다른 마이크를 통해 획득한 소리의 시간차를 이용하여 위상을 계산하고 칼만 필터를 사용하는 방법[3]과 고정된 마이크로 들어오는 소리의 입사각을 계산하고 밴드패스 필터를 사용하는 방법[4] 등이 있다. 그러나 사람은 두 귀에 도달하는 시간차이와 세기차이에 의해 소리의 방향을 인식하는 것으로 알려져 있다. 본 논문은 이에 착안 하여 그림 1과 같이 ITD (Interaural Time Difference)와 IID (Interaural Intensity difference)를 이용하여 음원의 방향을 인식한다.

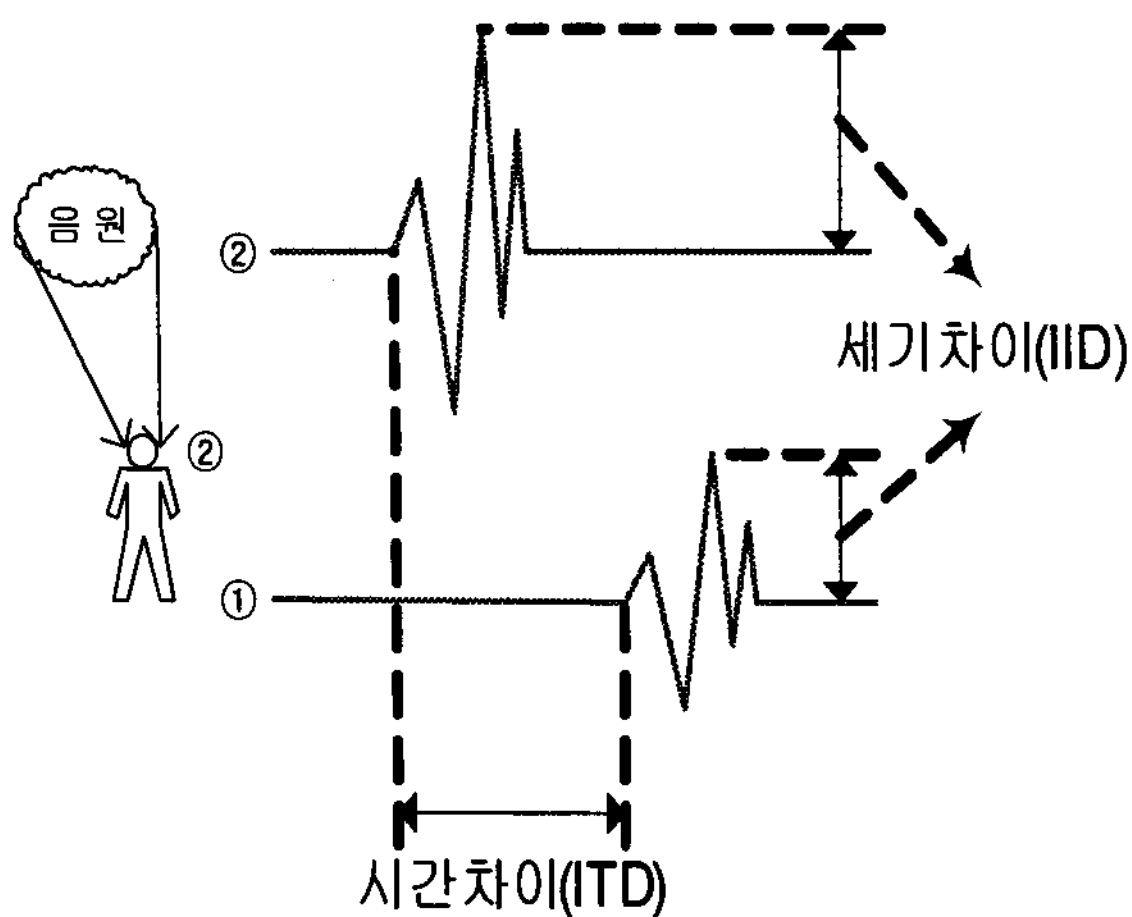


그림 1. 소리의 시간차이(ITD)와 세기차이(IID)

### 2.1 음원 추적센서의 구조

본 논문에서 제안하는 것은 음원의 정확한 위치를 검출하는 것이 아니라 음원이 발생한 쪽의 방향을 찾아내고, 대략적인 방향을 감지하는 것이다. 즉 음원의 대략적인 방향을 감지한 후 Pan-Tilt 카메라를 이용하여 물체를 추적하는 소스로 사용하는 것이다.

이를 위해서 그림 2와 같이 4개의 마이크를 카메라의 전면부에 설치한 뒤 음원이 발생 시 음원의 세기와 시간 차이를 계산하여 음원의 영역(A1~5)를 구분하여 음원의 대략적인 방향을 알아낸다.

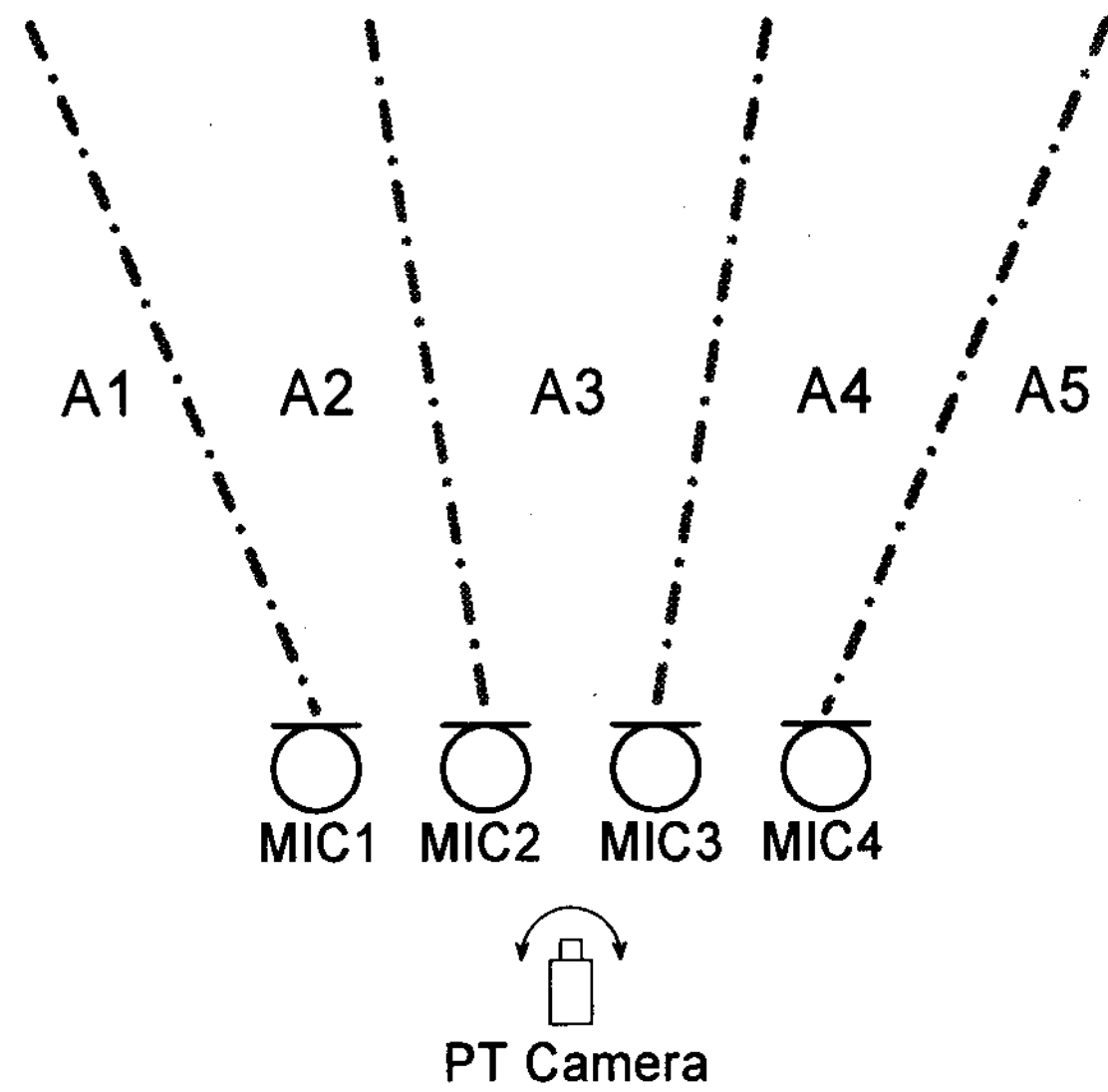


그림 2. 4 개의 마이크를 이용한 음원 위치 추정

## 3. PT 카메라를 이용한 움직임 추적

영상에서 객체의 움직임 정보를 추출 하는 방법 중 가장 널리 사용하는 방법은 차 영상을 이용하는 방법[5], 옵티컬 플로우(Optical flow)[6]와 블록 매칭(Block matching)[7]방법 등이 있다. 차 영상을 이용하는 방법은 기준 영상과 참조 영상의 밝기 차이를 구한 차 영상(difference image)을 기반으로 움직임 정보를 추출에 임계화 방법을 사용하기 때문에 임계치에 따라 잡음이 이동 물체로 검출 될 수도 있고, 이동 물체를 검출 하지 못할 수 있는 문제점이 있다. Optical flow는 과도한 계산 량을 요구하기 때문에 실시간 처리에는 적당하지 않다.

따라서 본 논문에서는 블록 매칭 움직임 추정 방법을 이용하여 영상 내 객체의 움직임 정보를 추출하고, Pan-Tilt 카메라를 이용하여 객체를 추적한다.

### 3.1 Pan-Tilt 카메라

영상 입력 장치로는 1/3", 41만 화소 CCD 컬러 Pan-Tilt 스펙의 소니 DSP 카메라를 사용하였다. Pan 312°, Tilt 80°이며, Pan-Tilt 제어는 485통신으로 제어한다. 프로토콜은 표 1과 같이 Synchronization 1byte, Command 2byte, Data 2byte, Checksum으로 구성되어 있는 PELCO사의 D로 제어한다.

카메라의 영상의 출력은 휘도 신호와 색차 신호가 함께 있는 복합 신호로 휘도와 색차 신호가 단일 반송파 신호에 함께 실리는 Composite로 출력된다.

표 1. PELCO-D 프로토콜

Byte1	Byte2	Byte3	Byte4	Byte5	Byte6	Byte7
Sync	Address	Cmd1	Cmd2	Data1	Data2	CRC

### 3.2 블록 매칭 움직임 추적 방법

이 방법은 평균 절대값 차이 (MAD : Mean Absolute Difference)를 이용하여 두 블록의 유사도를 평가하는 것이다. 즉  $t$ 번째 프레임  $I_t$ 에서 좌표  $(k,1)$ 에  $M \times N$ 크기의 블록이 있다고 가정했을 때 이 블록과  $t-1$ 번째 프레임  $I_{t-1}$ 에서 좌표  $(k+x,1+y)$ 에 위치한 블록과의 평균 절대 값 차이는 다음과 같이 계산한다[7].

$$MAD_{(k,y)}(x,y) = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} |I_t(k+i,l+j) - I_{t-1}(k+x+i,l+y+j)| \quad (1)$$

검색 윈도우 안에서 해당 블록과 MAD값이 가장 작은 위치의 블록을 찾고, 그 위치 변화를 움직임 벡터로 인지하게 된다.

## 4. 실험 결과 및 고찰

### 4.1 실험 장치의 구성

본 실험의 목적인 침입자 추적을 위해서 구현한 실험 장치는 그림 3과 같다. 우선 그림 2와 같은 음원 위치 추정을 위해 전면부에 음성을 받아들이기 위한 Pre Amp가 내장된 마이크가 4개 위치하며, 음성을 받아 들여 증폭하고, 디지털 값으로 변경하기 위한 회로가 구성 되어 있다. 또한 디지털 값으로 변경된 음성 데이터를 PC로 전송하기 위해 RS232와 USB로 서로 데이터를 변환해주는 FTDI사의 FT232를 사용 하였고, 컨트롤러로는 ATMEL사의 ATmega128을 사용하였다.

카메라의 영상은 Composite영상으로 출력되기 때문에 PC에서 입력 소스로 사용할 수가 없다. 따라서 PC로 받기 위해 USB 타입의 Video Capture Board를 사용하였으며, Pan-Tilt제어를 위한 PC와의 인터페이스를 위해 485통신을 USB로 변환하는 컨버터를 이용하였다.

### 4.2 침입자 추적 시스템의 구성

침입자 추적 시스템의 화면 구성은 그림 4와 같다. 카메라 입력은 실시간으로 카메라의 입력을 받아 표시하는 부분이다. 기준영상은 Pan-Tilt 카메라가 움직일 경우마다 기준영상

을 다시 획득하는데 이 영상을 표시한 것이다. 검출 결과는 카메라 입력과 기준영상의 차영상을 이진화하여 모폴로지 기법을 이용하여 잡음을 제거 한 화면이다. 또한 블록 매칭 움직임 추정 방법을 이용하여 움직임 벡터를 나타내었다. FT232는 물리적으로 USB인터페이스이지만 MS윈도우 에서는 가상 시리얼포트로 인식이 되기 때문에 Pan-Tilt 카메라제어 시리얼포트와 음원 추적 보드의 프로세서(ATMega128)와의 통신을 하기위한 시리얼 포트 설정 부분이 있다.

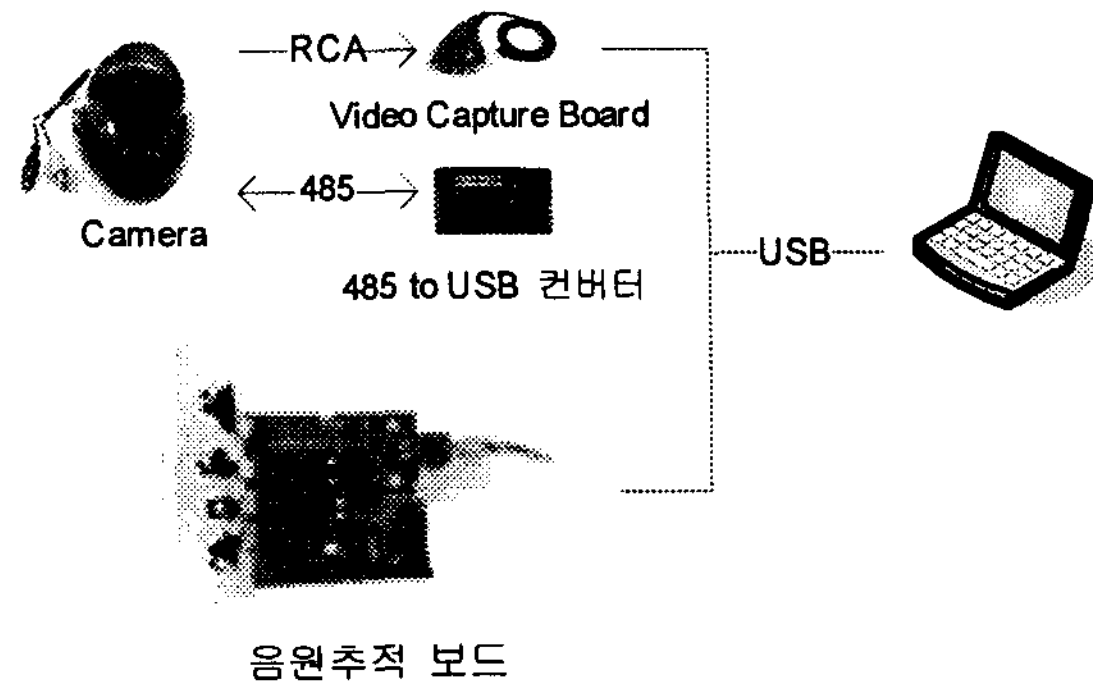


그림 3. 침입자 추적 시스템 구성도



그림 4. 침입자 추적 시스템 화면

### 4.3 실험 방법

음원의 위치 파악 및 침입자 추적 시스템의 추적 과정은 다음과 같다.

- [Step 1] 음원 추적 장치에서 마이크로 들어 오는 음원의 방향 벡터를 계산하고 위치 영역(A1-A5)을 결정한다. 만약 마이크로 음원이 들어오지 않을 경우에는 카메라 단독으로 객체의 움직임을 감지한다.
- [Step 2] 음원의 위치 영역을 결정 했거나 카메라 단독으로 움직임을 감지했을 경우에는

- 카메라를 이동시켜 움직임을 탐색한다.
- [Step 3] 움직임을 감지되었을 경우 감지 영상을 정지 영상으로 저장하고 감지되지 않았을 경우는 [Step 1]로 이동한다.
- [Step 4] 움직임을 감지된 정지 영상을 파일로 저장한 후 [Step 2]로 이동한다.

그림 5는 음원과 Pan-Tilt 카메라를 이용한 침입자 감지 시스템의 알고리즘을 나타낸다. 객체 추적 소스 음원을 사용하였지만 음원이 없을 경우는 카메라로 획득한 영상을 스스로 사용하였다.

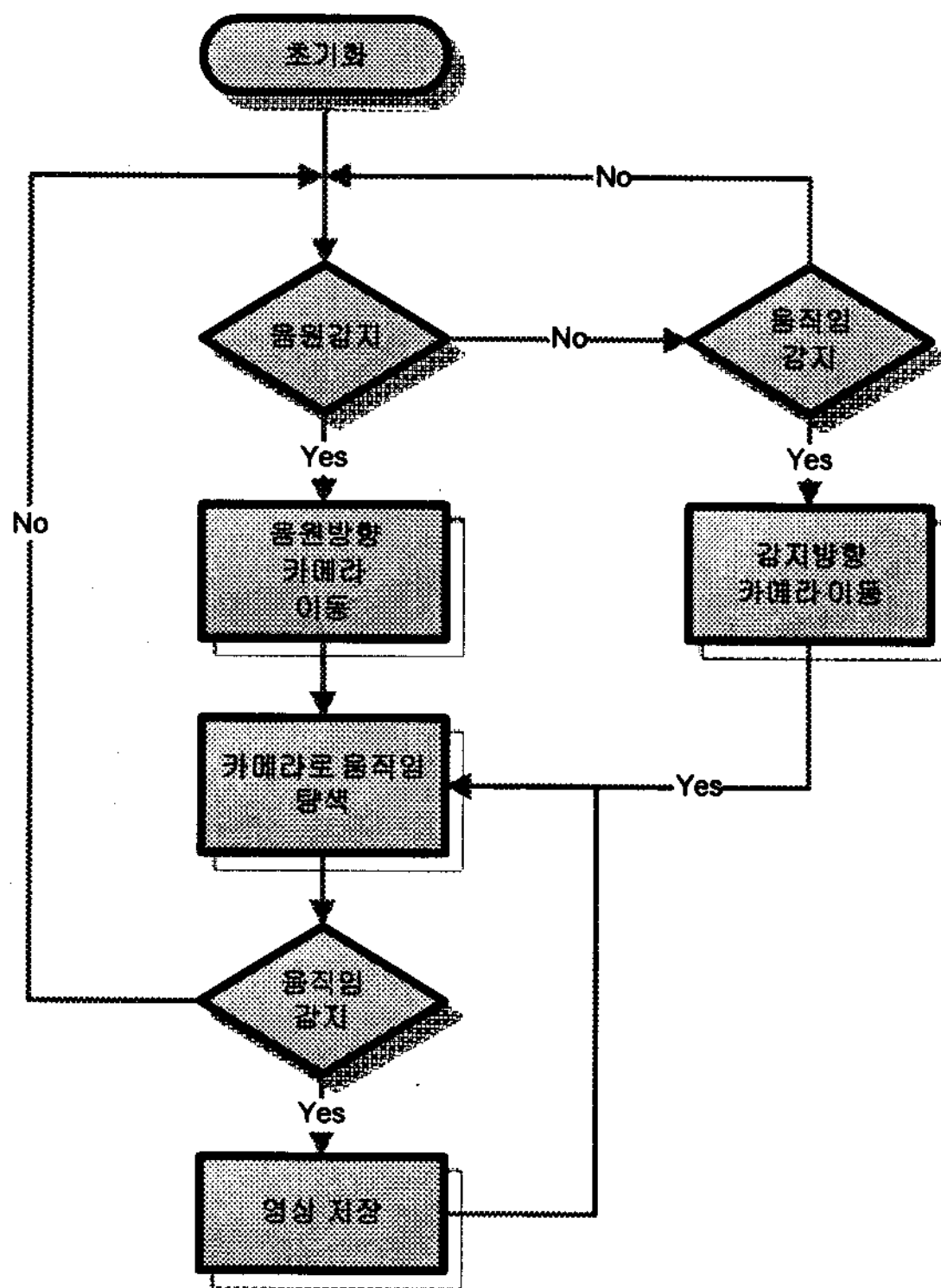


그림 5. 침입자 추적 과정

#### 4.4 실험 결과

시스템의 정확성을 알아보기 위하여 거리와 방향에 따른 음원 영역 구분 실험과 영상을 이용하여 객체를 추적하는 실험을 하였다. 음원의 위치 파악의 목적은 정확한 음원의 위치를 찾아내는 것이 아니었기 때문에 예상한 결과와 같았다. 그림 2의 영역구분에서 카메라와 거리가 멀어 질수록 영역이 바뀌어 나타나는 경우도 있었지만 카메라의 화각이 이를 커버할 수 있기 때문에 추적 시스템의 성능에는 크게 영향을 미치지 않았다. 그림 6은 추적 객체의 움직임 벡터를 표현한 것이다.

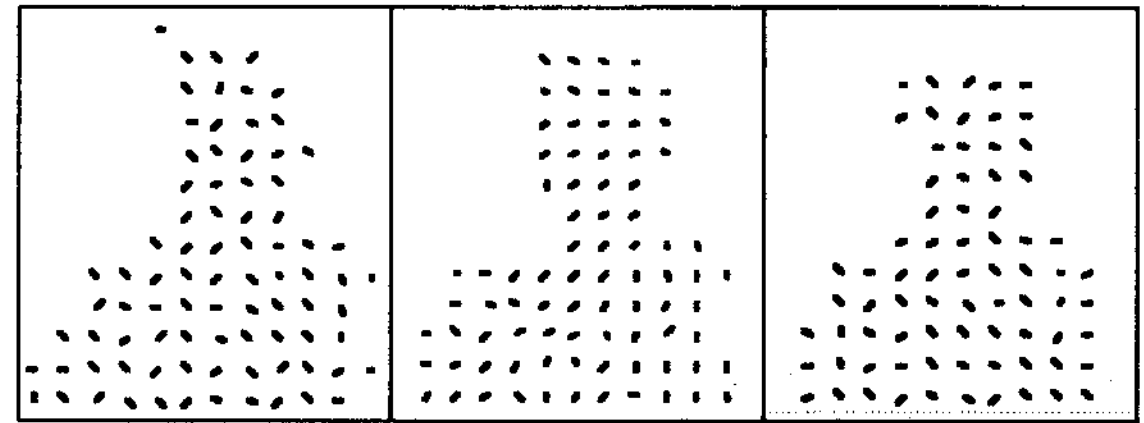


그림 6. 객체 추적 움직임 벡터

## 5. 결론 및 향후 과제

본 논문은 보안에 대한 중요성이 커짐에 따라 효과적인 보안 시스템구축의 필요성에 의해 음원의 위치 파악 및 침입자 추적 시스템을 구현 하였다. 시스템 구현에 있어서 침입자의 유효 영상을 얻기 위한 방법으로 Pan-Tilt 카메라를 사용했지만 침입자의 거리가 멀 경우에는 유효 영상을 얻지 못하였다. 이에 따라 카메라의 Pan-Tilt기능에 Zoom 기능이 포함된 PTZ 카메라를 이용하여 향후 효과적인 시스템 구현이 이루어져야 할 것이다.

## 참고 문헌

- [1] C.Micheloni, G.L. Foresti, L. Snidaro "A network of co-operative cameras for visual surveillance". *IEEE Proc. of Visual Image Signal Porcessing*. vol. 152, no.2, 2005.
- [2] A.W. Senior, A. Hampapur, M Lu "Acquiring Multi-Scale Images by Pan-Tilt-Zoom Control and Automatic Multi-Camera Calibration". *Proc. of Seventh IEEE Workshop on Applications of Computer Vision*, vol. 1, pp. 433-438, 2005.
- [3] 이지연, 한민수 "지능형 서비스 로봇을 위한 원거리 음원 추적 기술," *대한음성학회지*, 제57호, pp. 85-97, 2006.
- [4] 특허-특허권자:(주)조원정보, "감시시스템에서의 음원 탐지 및 제어 방법," 출원번호 10-2003-0010086.
- [5] Paul L. Rosin, "Thresholding for Change Detection," *Brunel University, technical report ISTR*, 1997.
- [6] J. L. Barron, et.all, "Systems and Experiment: Performance of Optical Flow Techniques", *Int'l J. of Computer Vision*, vol. 12, pp. 43-77, 1994.
- [7] 황선규, *영상처리 프로그래밍*, Hanbit Media, 2007.