

치안 향상을 위한 비명 감지 시스템

박기철*, 정인혜**, 권순환**, 형성재**, 김봉재**

*선문대학교 전자공학과

**선문대학교 컴퓨터공학과

e-mail : pkcjja147@gmail.com

Scream Detection System for Public Security Improvement

Kichoel Park*, Inhye Jeong**, Soonhwan Kwon**, Sungjae Hyung**, Bongjae Kim**

*Department of Electronic Engineering, Sun Moon University

**Division of Computer Science and Engineering, Sun Moon University

요 약

폐쇄회로 텔레비전(CCTV, Closed Circuit Television)은 범죄 예방을 위해서 현재 많이 사용되고 있다. 하지만, CCTV 의 경우 주로 영상에 기반하기 때문에 사각지대가 존재할 수 있으며 범죄 발생시 즉각적인 대처보다는 후속 조치에 주로 활용된다. 본 논문에서는 이와 같은 문제점을 해결하기 위해 비명을 탐지할 수 있는 치안 향상을 위한 비명 감지 시스템을 설계하고 구현하였다. 제안하는 시스템은 비명을 감지하면 경보음을 제공해주며, 중앙 시스템에 알림을 주어 관리자가 확인하고 즉각적인 대응이 가능할 수 있도록 할 수 있는 기능을 제공한다.

1. 서론

CCTV 는 영상과 같은 정보를 활용한 시각적인 방범 시스템으로 볼 수 있으며, 다양한 곳에서 방범용 CCTV 가 설치되어 활용되고 있다. 하지만 이러한 방범용 CCTV 에도 문제점이 있다. 전체 비중의 58.1% 정도는 고화질이 아닌 저화질이다. 저화질 방범용 CCTV 의 경우에는 야간에 기록된 영상 품질이 떨어져 그 활용도에 문제가 생기게 된다. 또한, 우범 지역임에도 불구하고 CCTV 가 제대로 설치되지 않은 곳이 대다수이다[1]. 이러한 실정으로 인해 사각지대를 비롯한 일부 지역에서는 범죄 예방이 어렵다는 문제가 있다[2].

한 연구의 결과를 살펴보면 방범용 CCTV 의 설치 대수가 많거나, 방범용 CCTV 가 설치된 지역일수록 범죄가 줄어들지는 않는 것으로 나왔다. 오히려 범죄가 증가하는 것으로 나타났다. 또한, 살인, 강간, 강도, 절도, 폭력 등 범죄 예방 효과를 살펴보면, 방범용 CCTV 의 설치 대수나 설치 여부는 유의한 영향을 미치지 못하는 것으로 나타났고, 방범용 CCTV 는 실제로는 범죄 예방의 효과가 미미한 것으로 나왔다[3]. 이와 같은 이유로 방범용 CCTV 는 범죄 예방의 목적보다는 범죄 발생 후 사건 해결에 더 유용하게 활용될 수 있다. 따라서, 방범용 CCTV 를 이용하여 범죄 현장을 즉각적으로 인지하고 반응하기는 쉽지 않다.

본 연구는 CCTV 의 시각적인 방범 시스템에 청각적인 요소인 비명 검출 기능을 추가한 치안 향상을 위한 비명 감지 시스템을 제안한다. 제안 시스템은 상황이 발생했을 때, 비명을 검출하여 신호를 보내줄 수 있고 CCTV 처럼 범죄에 대한 후속 조치를 위한

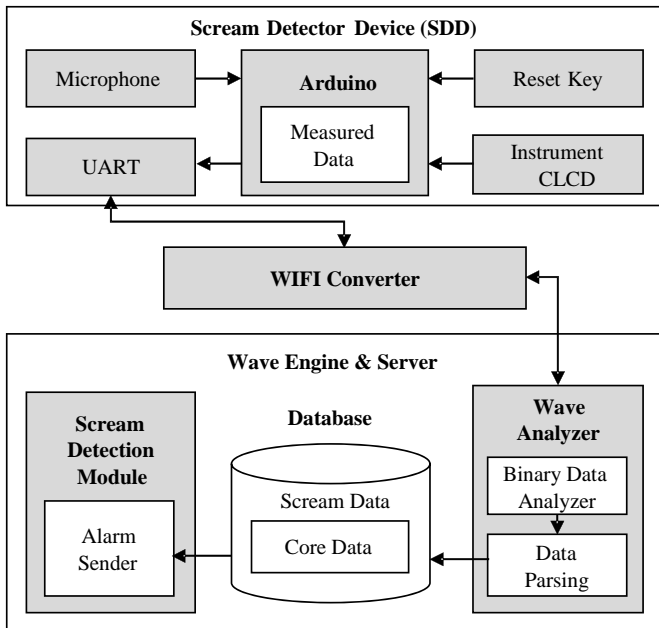
수단을 넘어 즉각적인 반응을 통하여 범죄를 예방하고 대응할 수 있다. 따라서 이와 같은 기능을 제공하여 기존 시각정보 위주인 방범용 CCTV 의 활용성과 기능성을 높아질 것 이라고 예상된다.

이후 본 논문의 구성은 다음과 같다. 2 장에서는 제안 시스템의 시스템 구조 및 서비스 흐름에 관해서 설명한다. 3 장에서는 제안 시스템의 설계 및 구현 내용에 대해서 자세히 알아본다. 4 장에서는 구현된 프로토타입 시스템의 구현 결과를 보이고, 마지막으로 5 장에서는 본 논문의 결론을 맺는다.

2. 시스템 구조 및 서비스 흐름

본 논문에서 제안한 비명 감지 시스템의 소프트웨어 도구는 Windows 7/.NET Framework 4.0 플랫폼 기반으로 설계하고 구현하였다. 자체 제작한 Wave Engine 을 통하여 입력된 음성을 분석하는 기능을 제공한다. 분석된 정보를 토대로 비명이 포함되어 있는지에 대한 여부를 판단하고, 그 결과를 전달할 수 있는 기능을 제공한다.

그림 1 은 제안한 비명 감지 시스템의 서비스 흐름을 보여주고 있다. SDD(Scream Detector Device)는 장착된 마이크를 통해 실시간 녹음 기능을 제공한다. 녹음된 음성 데이터는 비명 여부를 판단하기 위해 서버로 전송된다. 서버는 Wave 분석기(Wave Analyzer)를 이용하여 분석된 음성 정보를 데이터베이스와 비교한다. 비명 탐지 모듈(Scream Detection Module)은 데이터베이스에 있는 비명 정보와 대조하여 경보음을 이용해 탐지 여부를 전달한다.



(그림 1) 서비스 흐름도

3. 시스템 구조 및 서비스 흐름

3.1 WAVE 파일 개요

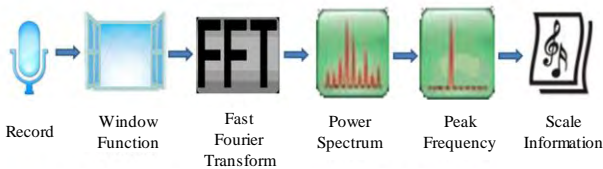
WAV 또는 WAVE 는 웨이브폼 오디오 포맷 (웨이브 오디오 포맷, Waveform Audio File Format)을 의미하며, 개인용 컴퓨터에서 오디오를 재생하는 마이크로소프트와 IBM 오디오 파일 포맷 표준이다. 가공되지 않은 오디오를 위한 기본 포맷이다. 그림 2 는 WAV 파일의 포맷의 상세내용을 보여준다. WAV 는 비압축 형식만 지원하는 것이 아니라 다양한 코덱을 이용하여 데이터에 대한 압축도 지원한다.

Chunk ID (4byte)	Chunk Size (4byte)	Format (4byte)	
Chunk ID (4byte)	Chunk Size (4byte)	Audio Format (2byte)	NumChannels (2byte)
Sample Rate (4byte)	Byte Rate (4byte)	Block Align (2byte)	Bits Per Sample (2byte)
Chunk ID (4byte)	Chunk Size (4byte)	(Low Data)	
(Low Data)			

- RIFF: Marks the file as a RIFF
- FMT : Format Chunk Marker
- DATA: Data Chunk

(그림 2) WAV 파일의 상세

3.2 음성인식



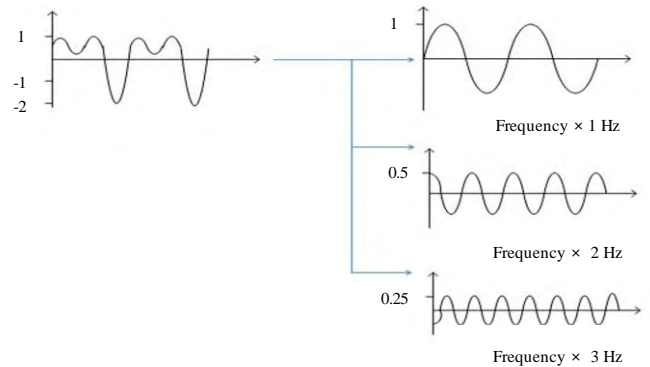
(그림 3) 음성인식 순서

그림 3 은 음성인식의 일반적인 순서를 보여준다. 그림 3 과 같이 음성 인식은 녹음된 음성 파일을 통해 임의의 일정 구간만을 가져와서 분석하는 창 함수 (Windows Function)를 사용한다. 창 함수를 통해 나온

데이터를 고속 푸리에 변환(FFT)을 사용하여, 특정 주파수를 찾아낼 수 있다. 주파수를 통해 우리가 쉽게 알 수 있는 음계 형태로 분석과 활용이 가능하다.

3.3 주파수영역에서의 특징 검출

WAV 는 소리 정보를 변형하지 않고 저장한 파일이다. 여기서 소리 정보는 무수히 많은 정형 파가 합쳐진 합성 파이다[4]. 그림 4 는 일반적인 합성 파 분석의 예를 보여준다. 그림 4 는 왼쪽의 파형이 일반적인 소리의 형태이다. 이러한 파형을 분석하여 일정 주기를 갖는 정형 파를 분리할 수 있다. 분리한 정형 파들의 가지고 있는 주파수를 음계의 주파수와 비교하여 음정을 인식 할 수 있다.



(그림 4) 합성 파 분석의 예

3.4 비명소리 검출

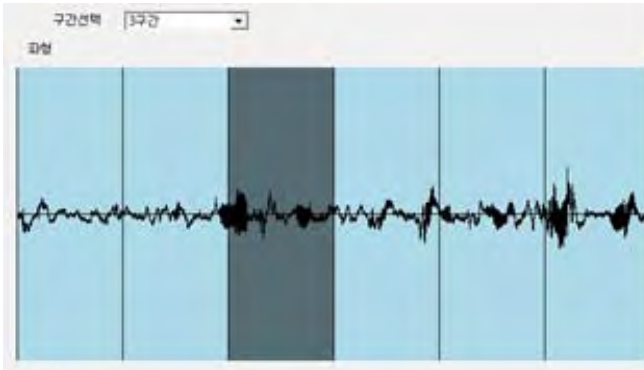
일반적인 도침에서 발생하는 소리는 공통적으로 300 Hz 이하이다. 비명과 혼동하기 쉬운 경적이나 노랫소리 같은 날카로운 소리는 2,500 Hz 부근에서 큰 에너지가 발생한다. 변화가 환경에서 주로 발생하는 소리는 사람들의 말소리와 노래소리이며, 주로 2,000 Hz 부근에서 큰 에너지를 보인다 [5].

<표 1> 주파수에 따른 표준음역대

음역대(Hz)	1	2	3	4	5	6	7	8
C(도)	32.70	65.41	130.81	261.63	523.25	1046.50	2093.01	4186.01
C#	34.65	69.30	138.60	277.18	554.37	1108.73	2217.46	4434.92
D(레)	36.71	73.42	146.83	293.66	587.33	1174.66	2349.32	4698.64
D#	38.89	77.78	155.56	311.13	622.25	1244.51	2489.02	4978.03
E(미)	41.20	82.41	164.81	329.63	659.26	1318.51	2637.02	5274.04
F(파)	43.65	87.31	174.61	349.23	698.46	1398.91	2793.83	5587.65
F#	46.25	92.50	185.00	369.99	739.99	1479.98	2959.96	5919.91
G(솔)	49.00	98.00	196.00	392.00	783.99	1567.98	3135.96	6271.93
G#	51.91	103.83	207.65	415.30	830.61	1661.22	3322.44	6644.86
A(라)	55.00	110.00	220.00	440.00	880.00	1760.00	3520.00	7040.00
A#	58.27	116.54	233.08	466.16	932.33	1864.66	3729.31	7458.62
B(시)	61.74	123.47	246.94	493.88	987.77	1975.53	3951.07	7902.13

음역이 낮은 사람의 경우 비명 주파수 대역은 150Hz 에서 533Hz 사이에 나타나고, 높은 사람의 경우 비명의 주파수는 500Hz에서 2133Hz에서 나타난다 [6]. 남성의 비명의 경우는 500Hz 에서 1,500Hz 부근

대역에 특징적인 에너지가 나타나고, 여성의 경우는 1,000Hz 에서 2,000Hz 부근 대역에서 특징적인 에너지를 나타낸다[7]. 이러한 정보를 토대로 비명에 대해 구분할 수 있다. 그림 5 는 대화 소리에 대한 주파수 영역 그래프의 예를 보여준다.



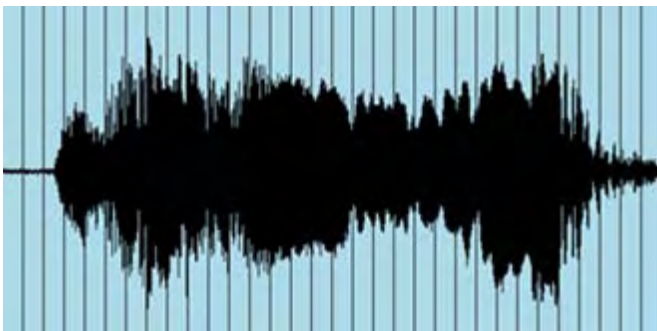
(그림 5) 대화 소리에 대한 파형의 예

3.5 비명 구간 찾기

비명 구간을 찾기 위해 짧은 시간 동안 녹음을 하고 분석을 하여 비명에 대한 패턴을 찾아내는 방식으로 하였다. 비명의 시작점을 찾고, 조건에 만족하는 프레임이 일정 구간 이상이 연속하였을 때, 비명으로 인식한다.

4. 실험 결과

본 논문에서 제안된 비명 감지 시스템의 성능의 실험에서는 일반적인 환경 잡음이 있는 상태에서 1 m 정도의 거리에서 음성을 획득하는 형태로 실험을 진행하였다. 획득한 음성 데이터는 비명 여부인지를 분석하고 판단하기 위해 0.1초 간격 분할하였다. 그림 6 은 실험에서 사용된 여성 비명소리 실험 데이터의 파형의 예이다. 표 2 는 그림 6 의 여성비명소리의 탐지 결과에 대한 상세내용이다.



(그림 6) 여성 비명소리 실험 데이터의 파형의 예

<표 2> 그림 6 의 여성비명소리의 탐지 결과

분석 영역 번호	주파수(Hz)	음정	탐지 여부
1	329.63	41	X
2	6644.86	93	X
3	3135.96	80	O

4	3729.31	83	O
5	5919.91	91	X
6	5919.91	91	X
7	2793.83	78	O
8	2793.83	78	O
9	5274.04	89	X
10	2489.02	76	O
11	2489.02	76	O
12	2489.02	76	O
13	2489.02	76	O
14	2489.02	76	O
15	2489.02	76	O

표 2 는 여성의 비명소리 중 15 구간의 데이터만 작성했다. 음정은 표 1 에 해당하는 주파수영역을 표현했다.

구현된 비명 감지 시스템은 총 50 개의 분석 영역에서 여성 비명에 해당하는 영역 30 개 중 26 개를 탐지하였으며, 탐지율은 86.6 % 이다.

5. 결론 및 향후 연구

기존의 시각적인 정보에만 의지하는 방법용 CCTV 의 경우 범죄에 대한 즉각적인 대응이 어렵다. 또한 한 사람이 여러 개의 영상을 감시하게 되면 시간이 지날수록 그 효율성이 떨어지는 문제도 발생할 수 있다. 본 연구는 이와 같은 문제들을 해결할 수 있는 치안 향상을 위한 비명 감지 시스템을 제안하였다. 제안 기법은 비명을 탐지하고, 이를 경보음 형태의 정보로 제공하여 즉각적인 대응이 가능할 수 있도록 지원한다. 구현된 프로토타입의 실험을 통해 비명을 효과적으로 검출할 수 있음을 검증하였다. 따라서 기존의 방법 시스템을 효율성과 기능성을 높일 수 있으며, 범죄의 발생률과 범인의 검거 시간의 단축에도 기여할 수 있을 것이다.

향후 연구에서는 생활 잡음을 배제하고 탐지율을 개선하고, 음성 정보뿐만 아니라 딥러닝과 같은 기술을 영상 정보 분석에 같이 적용하여 오탐지율을 개선하는 연구를 진행하고자 한다.

사사

이 성과는 2017 년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2017R1C1B5017476)

참고문헌

- [1] 박호정. “방법용 CCTV 의 검거효과에 관한 연구.” 한국보안 논문지 14.6 (2014): 25-33.
- [2] Haritaoglu, Ismail, David Harwood, and Larry S. Davis. “W/sup 4: real-time surveillance of people and their

- activities.” IEEE Transactions on pattern analysis and machine intelligence 22.8 (2000): 809-830.
- [3] 임민혁, “방법용 CCTV 를 활용한 범죄예방 효과 연구”, 중앙대학교 대학원(2007), 석사학위 논문
- [4] 박철원, 신명철. “특정 주파수 추출을 위한 여러 가지 디지털 필터의 특성 비교.” 대한전기학회 학술대회 논문집 (2004): 360-362.
- [5] 박주현, et al. “오디오 취득 기반의 방법용 시스템을 위한 환경 잡음 분석.” 한국방송공학회 학술발표대회 논문집 (2013): 81-84.
- [6] Chan, Cheung-Fat, and W. M. Eric. “An abnormal sound detection and classification system for surveillance applications.” Signal Processing Conference, 2010 18th European. IEEE, 2010.
- [7] Lee, So-Min, et al. “Screaming data analysis for security system with audio capability.” presented at the Korean Society of Broadcast Engineers autumn Conference. 2013.