MCLT 피크쌍 기반의 강인한 해시 함수를 이용한 오디오 핑거프린팅

Audio Fingerprinting Using a Robust Hash Function Based on the MCLT Peak-Pair

이준용, 김형국

(Jun-Yong Lee and Hyoung-Gook Kim[†])

광운대학교 전파공학과

(Received November 18, 2014; accepted January 16, 2015)

초 록: 본 논문은 MCLT(Modulated complex lapped transform) 피크쌍 기반의 강인한 해시 함수를 이용한 오디오 핑거프린팅을 제안한다. 기존 방식의 오디오 핑거프린팅은 시간-스케일, 피치-이동, 이퀄라이제이션과 같은 다양한 왜곡이 발생했을 때, 강인한 핑거프린트를 추출하지 못한다는 문제점이 있다. 이를 해결하기 위해 본 논문에서는 MCLT 스펙트럼, 현저한 피크검색을 위한 적응적 문턱값 방식, 개선된 해시 함수를 이용하여 잡음과 왜곡 환경에 강인한 오디오 핑거프린팅을 추출하였다. 실험결과 제안된 방식이 잡음과 왜곡 환경에서 보다 강인한 핑거프린팅을 추출할 수 있으며 매칭 인식률을 향상 시키는 것을 보여준다.

핵심용어: 오디오 핑거프린팅, MCLT, 피크 탐지, 강인한 해시 함수

ABSTRACT: In this paper, we propose an audio fingerprinting using robust hash based on the MCLT (Modulated Complex Lapped Transform) peak-pair. In existing methods, the robust audio fingerprinting is not generated if various distortions occurred; time-scaling, pith-shifting and equalization. To solve this problem, we used the spectrum of the MCLT, an adaptive thresholding method for detection of prominent peaks and the novel hash function in the audio fingerprinting. Experimental results show that the proposed method is highly robust in various distorted environments and achieves better identification rates compared to other methods.

Keywords: Audio fingerprinting, Modulated complex lapped transform, Peak detection, Robust hash function **PACS numbers:** 43.60.-c

I. 서 론

최근 모바일 단말기에서 음원검색,^[1] 복제 음원 검색, 광고 검색, 등 다양한 애플리케이션이 개발되어 사용되고 있다. 이러한 애플리케이션에 적용된 오디오 핑거프린팅 기술은 단구간의 질의 오디오 클립을 빠르고 정확하게 식별해 주는 방식으로써 이 기술이 적용되기 위해서는 다음과 같은 조건들이 충족되어 야한다: (1) 시간-스케일, 피치 이동, 이퀄라이제이

선, 다양한 잡음과 아티팩트에 강인해야 한다. (2) 단구간의 오디오 클립에 대해서 음악 검색이 가능해야한다. (3) 저 연산의 효율적인 방식으로 핑거프린팅생성과 음악검색을 수행해야 한다.

pISSN: 1225-4428 eISSN: 2287-3775

위 조건을 충족하기 위해 다양한 오디오 핑거프린 팅 방식들이 개발되었는데 그중 Wang^[2]이 제안한 방 식이 가장 널리 사용되고 있다. 이 방식은 효율적인 오디오 핑거프린트 산출과 대용량의 데이터베이스 에서 음악검색을 수행할 수 있다는 두 가지 조건을 충족시킨다. 하지만 Wang^[2]의 방식은 오디오 신호의 피치 이동 혹은 시간-스트레치와 같은 왜곡이 발생 했을 때 오디오 핑거프린트를 통한 검색정확도가 현

†Corresponding author: Hyoung-Gook Kim (hkim@kw.ac.kr) Department of Wireless Communication Engineering, Kwang-Woon University, 20 Gwangun-Ro, Nowon-Gu, Seoul 139-701, Republic of Korea

(Tel: 82-2-940-5574, Fax: 82-2-940-5110)

저하게 저하되는 문제점이 있다.

따라서 이러한 문제를 해결하기 위해 본 논문에서는 Wang^[2]의 방식을 기반으로 개선된 오디오 핑거프린팅 방식을 제안한다. 제안된 방식은 다음과 같은 4 가지의 이점이 있다: (1) MCLT(Modulated Complex Lapped Transform) 기반의스펙트럴 피크 추출은 STFT (Short Time Fourier Transform) 기반스펙트럴 피크 추출보다 음원 본연의 피크를 더 효과적으로 추출하여 검색정확도를 향상시킨다. (2) 메디언 필터 기반의 적응적 문턱값 방식과 피크 추출 갱신 방법을 적용하여 다양한 왜곡 환경으로부터 강인한 피크를 추출한다. (3) 개선된 핑거프린트 해시를 사용함으로써잡음, 피치 이동, 시간-스케일에 강인한 오디오 핑거프린트를 검색에 효과적으로 적용한다. (4) 저연산, 고성능의 검색 결과를 제공할 뿐 아니라 모바일 단말기와 호환성이 뛰어나다.

본 논문은 다음과 같이 구성되어있다. II 장에서는 제안된 방식에 대해 기술하고 III 장에서는 실험 결과를 제시하며 마지막으로 IV장에서 결론 및 향후 연구 방향을 서술한다.

비. 강인한 오디오 핑거프린트 추출 방식

강인한 오디오 핑거프린트 추출방식은 Fig. 1과 같이 5개의 모듈로 구성되어 있다.

먼저 사용자의 모바일 단말기로부터 스테레오 오디오 신호가 녹취되고 이 신호는 16 kHz 샘플링레이트의 모노 신호로 변환된다. 변환된 신호는 해밍 윈도우 함수를 적용하여 오버래핑 되며 스펙트럴 피크 검출을 위해 로그 MCLT 스펙트럼(LMCLT) 이 각 프레임에 적용된다. 이때 로그 MCLT 스펙트럼에 메디언 필터가 적용되어 고에너지의 값을 가지는 부분만을 현저한 스펙트럴 피크로 추출하고 이 피크들은 주파수 시간 축의 지역적인 설정 범위 안에서 피크 쌍으로 결합한다. 결합된 피크쌍은 사용자 단말기에서 32 비트 핑거프린트 해시로 변환된 후, 콘텐츠 식별을 위해 핑거프린트 서버로 전송된다.

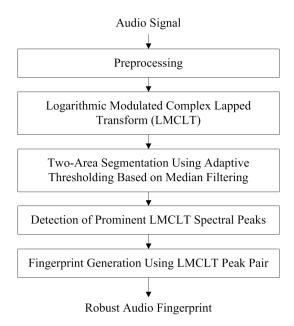


Fig. 1 Block diagram for robust fingerprint extraction.

2.1 Modulated Complex Lapped Transform

먼저 오디오 신호s(n)이 해밍 윈도우가 적용된 중복 프레임으로 분할되고 MCLT가 적용된다. FFT와 같은 직교 변환과 달리 MCLT는 주파수 해상도의 2배만큼의 주파수 응답 중복을 가진다. 따라서 MCLT는 시프트 변화에 거의 영향을 받지 않는다. [3] MCLT 기반의 스펙트럴 피크 추출보다 음원 본연의 피크를 더 효과적으로 추출하는데 특히 잡음, 에코, 아트팩트와 같은 다양한 왜곡에 강인한 특성을 보인다. 왜곡에 강인한 피크 추출은 짧은 쿼리의 오디오 클립이 입력될 때 보다 더 높은 정확도로 오디오 트랙을 식별할 수 있도록 한다. 다음 식은 MCLT 스펙트럼 $S_{MCLT}(k,l)$ 을 계산하는 과정을 나타낸다.

$$\begin{split} S_{MCLT}(k,l) &= |i\,V(k,l) + V(k+1,l)| \\ V(k,l) &= b(k,l)\,U(k,l) \\ U(k,l) &= \sqrt{\frac{1}{2N}} \sum_{n=0}^{2N-1} s\,(n+lM)h(n) \mathrm{exp}(\frac{-i2\pi kn}{N}) \,, \\ b(k,l) &= W_8(2k+1,l)\,W_{4N}(k,l) \\ W_T(k,l) &= \mathrm{exp}(\frac{-i2\pi k}{T}) \end{split} \tag{1}$$

여기서 k는 주파수 빈, l은 프레임 인덱스이며 n은 입

력된 오디오 신호의 샘플 인덱스이다. h는 N크기의 윈 도우 함수를 나타내며 M은 프레임 스텝을 나타낸다.

2.2 로그 MCLT 스펙트럼 계산과 메디언 필터 기반의 강인한 피크 추출

로그 함수는 MCLT 스펙트럼 $S_{MCLT}(k,l)$ 을 가청 크기로 나타내기 위해 사용된다. 로그 MCLT 스펙트럼 계산 과정은 다음과 같다.

$$S_{log}(k,l) = log_{10}[S_{MCLT}(k,l)].$$
 (2)

계산된 로그 MCLT 스펙트럼 $S_{\log}(k,l)$ 으로부터강인한 피크를 추출하기 위해 2차원의 메디언 필터를 적용하는데 이 방식은 잡음에 민감한 고주파 대역의스펙트럴 피크를 추출할 수 있는 장점이 있다.

$$M(k,l) = \underset{(r,s) \in w}{median} S_{\log}(k+r,l+s), \tag{3}$$

여기서 w는 메디언 필터가 적용된 윈도우의 크기를 나타낸다.

메디언 필터 기반의 로그 MCLT 스펙트럼 M(k,l)을 계산하기 위해 적용되는 메디언 필터 범위는 10 < r < 40, 5 < s < 20로 구성되어 있으며 이 값은 실험을 통해 도출된 최적의 필터 범위이다. 계산된 메디언 필터 기반의 로그 MCLT 스펙트럼과 로그 MCLT 스펙트럼의 값을 서로 비교하여 고 에너지를 갖는 로그 MCLT 스펙트럼 $S_{LM}(k,l)$ 을 계산한다.

$$S_{LM}(k,l) = \begin{cases} S_{\log}(k,l), & \text{if } S_{\log}(k,l) > M(k,l) \\ 0, & \text{otherwise} \end{cases}$$
 (4)

위 과정을 통해 산출된 고 에너지를 갖는 로그 MCLT 스펙트럼 값은 주변 지점에서 가장 강한 에너지를 가 지고 있기 때문에 강인한 피크로써 추출된다.

2.3 적응적 문턱값 기반의 고 에너지 로그 MCLT 스펙트럴 피크 검색

적응적 문턱값 기반의 현저한 피크 추출 방식은 순방향 피크 검출, 역방향 피크 검출 두 가지 방식으 로 구성되어 있다. 세부 수행단계는 다음과 같다.

(단계 1) 초기 문턱값 선정 : $S_{LM}(k,l)$ 의 첫 번째 프레임 (l=0)에서 각 주파수 빈값들이 선형으로 보간되고 이 값들은 순방향 피크 검출 방식의 초기 문턱 값으로 사용된다.

(단계 2) 초기 문턱값 설정 이후 순방향 피크 검출 : $S_{LM}(k,l)$ 의두 번째 프레임 (l=1)에서 초기 문턱값과 $S_{LM}(k,l)$ 값을 비교하여 $S_{LM}(k,l)$ 값이 더 크다면 현재 프레임의 모든 피크들을 순방향 피크 $P_f(k,l)$ 의 튜플 (k,l)로 저장한다.

$$P_f(k,l) = \begin{cases} \text{forward peak} & \text{if } S_{LM}(k,l) > Th_F(k,l) \\ \text{non-forward peak} & \textit{otherwise} \end{cases} . \quad \text{(5)}$$

두 번째 프레임 이후 순방향 문턱값 $Th_F(k,l)$ 을 이용해서 순방향 피크를 추출한다. 다음 Eq. (6)은 순방향 문턱값 설정 방법을 나타낸다.

$$\begin{split} Th_F(k,l) &= \alpha \quad \bullet \quad Th_{\mathit{UP}}(k,l-1) + (1-\alpha) \\ & \quad \bullet \quad P_f(k,l-1)^* \exp\biggl(-\frac{m^2}{2\sigma^2}\biggr), \end{split} \tag{6}$$

$$Th_{\mathit{UP}}(k,l) = \max \left[Th_F(k,l-1), P_f(k,l)^* \exp \left(-\frac{m^2}{2\sigma^2} \right) \right], \eqno(7)$$

여기서 m은 이전 프레임의 피크와 현재 프레임에서 추출된 피크간의 거리이고 α는 스무딩 파라미터,σ 는 문턱값의 앤밸롭 첨도를 결정하는 파라미터이다. 그리고 Th_{UP}(k,l)은 이전 프레임에서 추출된 피크 값이 다음 프레임에서 추출되는 것을 방지하기 위해 사용되는 갱신된 순방향 문턱값을 나타낸다. 갱신된 순방향 문턱값을 통해 현재 프레임에서 검출된 피크의 상승분만큼 갱신되는 적응적 문턱값을 생성 할수 있다.

(단계 3) 역방향의 현저한 피크 검색: 역방향 피크 추출은 순방향 피크 검출을 검증하는 단계로써 왜곡 과 잡음에 더 강인한 피크 쌍을 추출해 주는 역할을 한다.

역방향 피크 추출을 하기 위해서는 순방향 피크 추출 이후 마지막 프레임 값을 역방향 피크 추출의 문턱값 $Th_B(k,l)$ 으로 설정한다. 그리고 순방향 피

160 이준용, 김형국

크에서 저장된 튜플 (k,l) 값들과 비교하여 역방향의 현저한 피크 $P_B(k,l)$ 만을 추출한다.

$$P_{B}(k,l) = \begin{cases} \text{prominent peak} & \text{if } P_{f}(k,l) > \textit{Th}_{B}(k,l) \\ \text{non-prominent peak} & \textit{otherwise} \end{cases}. \tag{8}$$

다음 Eq. (9)은 역방향 문턱값 설정 방법을 나타낸다.

$$\mathit{Th}_B(k,l) = \max \left[\mathit{Th}_B(k,l-1), P_B(k,l) * \exp \left(-\frac{m^2}{2\sigma^2} \right) \right]. \eqno(9)$$

2.4 로그 MCLT 스팩트럴 피크 쌍을 이용한 핑거프린트 해시

Wang^[2]의 방식에서 핑거프린트 해시는 랜드마크 지점이라 명명된 피크 쌍으로 구성된다. 각 피크쌍 은 목표영역 안에 있는 기준 피크를 기준으로 인접 한 피크를 선정하는데 이때 목표영역은 일정 수의 피크 쌍만을 생성하기 위해 사용된다.

목표영역 내에서 선정된 기준 피크 $P_B(k_a, l_a)$ 와 인접 피크 $P_B(k_a, l_a)$ 로부터 해시는 다음과 같이 생성된다.

$$hash = [k_a, k_a - k_p, (l_a - l_p)] = (k_a, \Delta k, \Delta l),$$
 (10)

여기서 모든 k (주파수빈) 와 l (프레임)은 고정된 바운드와 정수이다.

Wang^[2]의 방식에서 사용된 해시는 잡음에 매우 강 인하고 큰 데이터베이스에서도 높은 인식률을 보이 지만 오디오 신호의 피치 이동 혹은 시간-스트레치 에 취약하다는 문제점이 있다. 따라서 이 문제를 해 결하기 위해 다음과 같은 개선된 해시 방법을 제안 하다.

$$\begin{aligned} hash &= \left[\varepsilon \bullet \frac{k_a}{k_p}, \theta \bullet \frac{k_a - k_p}{k_a}, (l_a - l_p) + (k_a - k_p) \right] \\ &= \left(\varepsilon \bullet k_R, \theta \bullet \frac{\Delta k}{k_a}, \Delta l + \Delta k \right) \end{aligned} , \quad (11)$$

여기서 첫 번째 변수 k_R 는 피크 쌍의 시작 주파수 지점에 대한 끝 주파수 지점의 비율을 나타낸다. 그리고 두 번째 변수인 $\Delta k/k_a$ 는 피크 쌍 간의 주파수 길이의 차이에 대한 시작 주파수 지점의 비율을 나타낸

다. 첫 번째와 두 번째 변수는 피치 이동과 시간-스트 레치에 대해 강인함을 보여주지만 비율 기반의 해시 값은 매우 작은 범위에서 형성되는 문제가 발생한다. 따라서 해시 값에 ϵ 와 θ 의 가중치를 적용한다.

2.5 핑거프린트 매칭

서버에서는 핑거프린트 데이터베이스와 핑거프 린트 해시에 해당하는 인덱스와 트랙 ID가 설정된 다. 생성된 인덱스와 트랙 ID를 통해 다음과 같은 과 정을 거쳐 매칭 과정이 수행된다.

(단계 1) 사용자의 단말기로부터 입력된 쿼리 신호의 핑거프린트가 추출되고 서버로 전송된다. 쿼리파일의 해시는 서버의 해시테이블에 저장된 데이터베이스의 해시 값과 비교 후 매칭된다.

(단계 2) 매칭 된 데이터베이스 해시 값들로부터 음원 \mathbb{ID} S_{ID} , 시간의 길이 t_R , 주파수의 길이 f_R 를 각각 검출하고 같은 음원의 \mathbb{ID} 로 매칭된 S_{ID} 의수를 센다. 그리고 내림차순으로 정리하여 생성된 음원 \mathbb{ID} 리스트로부터 상위 30 %의 음원 \mathbb{ID} 만을 매칭결과 후보 \mathbb{ID} 로 선정한다.

(단계 3) 선정된 후보 $ID로부터 데이터베이스의 핑거프린트의 시간의 길이인 <math>t_R$ 과 쿼리 핑거프린트의 시간의 길이 t_Q 를 차감하여 핑거프린트 쌍 간의 시간 차 $\{t_R-t_Q\}$ 와 주파수 차 $\{f_R-f_Q\}$ 를 계산한다. 계산된 값들은 히스토그램에 저장되는데 만약질의된 오디오 프레임과 데이터베이스 안에 저장된 프레임이 동일한 핑거프린트라면 히스토그램의 값은 작게 산출될 것이다.

(단계 4) 잡음, 피치-이동, 시간-스트레칭 3가지의 단계로 나누어서 최종 매칭 과정이 다음과 같이 수 행되다.

- 쿼리 파일이 환경 잡음에 의해 왜곡되었다면 쿼리 파일의 시간 차와 주파수 차의 합은 기정된 문턱 값(가장 낮은 히스트로그램 범위) 안에 형성될 것이다. 형성된 값을 통해 검색된 음원 ID를 매칭 결과로 사용한다.
- 쿼리 파일의 피치-이동이 발생했다면, 주파수의 차이는 기정 시간 문턱값 보다 높게 형성되고 시간 차이는 기정 주파수 문턱값 보다 낮게 형성된다. 이 경우 히스토그램 기반의 주파수 차는 사용하지 않고

시간 차의 히스토그램만 사용하여 매칭 결과를 나타 내다.

- 쿼리 파일이 시간-스트레치가 발생했다면, 시간 차는 기정 시간 문턱값 보다 높게 형성되고 주파수 차는 기정 주파수 문턱값 보다 낮게 형성된다. 이 경우 히스토그램 기반시간 차는 사용하지 않고 주파수 차의 히스토그램만 사용하여 매칭 결과를 나타낸다.

마지막으로 매칭 과정의 연속성, 중복성을 피하기 위해 통계 필터를 적용한다. 이 과정은 쿼리와 데이 터베이스의 매칭 정확도를 향상시킨다.

Ⅲ. 실험 결과

3.1 실험데이터 구성

실험을 위해 팝, 락, 힙합, 포크, 재즈, 클래식의 9000개 음원의 데이터베이스를 구성하였다. 모든 오디오 데이터는 모노의 PCM 파일로 저장되어있으며 모바일 단말기와 같은 휴대장치에서 사용할 수 있도록 16 비트의 해상도, 16 kHz의 샘플링레이트로 구성하였다. 다양한 환경의 왜곡을 부여하기 위해 질의 (쿼리) 오디오데이터를 다음과 같이 구성했다: (1)잡음: 신호 대 잡음비 (SNR) 10 dB, 5 dB, 0 dB의 다양한 형태의 잡음(배블 잡음, 움직이는 자동차 잡음, 백색 잡음, 길거리 잡음, 컴퓨터 잡음), (2) 이퀄라이제이션: 31 Hz에서 16 kHz까지 이득 값-5 dB, 3 dB, (3) 시간스트레치: -20%부터 +20%까지 시간 변화, (3) 피치 이동: -30 %부터 +30 %까지 피치 변화 그리고 쿼리 오디오클립은 3 s, 4 s, 5 s로 모바일 단말기에서 녹취되었다. [4]

3.2 실험 결과

본 논문에서 제안한 방식을 비교하기 위해, 기존에 잘 알려진 세 가지의 방식이 사용되었고 각 방식의 파라미터 값들은 좋은 인식 결과를 얻기 위해 최적화를 시켰다. 방식 1은 본 논문에서 제안된 방식이고, 방식 2는 Wang^[2]의 STFT 방식의 피크 쌍 핑거프린트 추출 방식, 방식 3은 오디오 핑거프린팅의 두 방향성 피크 탐지 방식,^[4] 그리고 방식 4는 Constant Q 결합 방식의 해싱 기술을 적용한 핑거프린팅 추출 방식이다.^[5]

Table 1. Comparative performance for the four schemes.

SNR	Averaged Identification Rate (%)					
	Method 1	Method 2	Method 3	Method 4		
Clean	96.3	93.1	91.6	95.1		
10 dB	95.2	90.4	77.2	94.3		
5 dB	91.9	81.5	62.3	85.6		
0 dB	80.3	68.8	58.8	72.4		
Total	90.9	83.5	72.5	86.9		

Table 2. Performance evaluation according to query length.

SNR	Averaged identification rate (%)				
SINK	3 s	4 s	5 s		
Clean	92.6	96.3	98.1		
12 dB	91.8	95.2	97.7		
6 dB	86.2	91.9	95.4		
0 dB	78.4	80.3	82.6		
Total	87.3	90.9	93.5		

Table 3. Comparative performance of four schemes.

Type of distortion	Averaged identification rate (%) using 4 s query					
distortion	Method 1	Method 2	Method 3	Method 4		
EQ	96.3	92.8	90.5	95.1		
NA	90.9	83.5	72.5	86.9		
PS	87.4	23.7	22.3	52.3		
TS	86.3	22.5	21.5	40.1		
Total	92.7	71.6	68.7	79.9		

Table 1은4s 쿼리 길이에 따른5가지 다른잡음 환경에서의 각 방식에 대한 비교 실험 결과를 나타낸다.

인식률은 데이터베이스의 원 음원과 쿼리 간의 정확하게 매칭 된 검색 결과를 나타낸다. 실험 결과 본논문에서 제안한 방식 1의 클린 환경에서의 결과가 96.3 %로 가장 높은 인식률을 보였다.

Table 2는 방식 1의 쿼리 길이의 변화에 따른 각 잡음별 인식률을 나타낸다. 실험 결과 쿼리의 길이가 늘어날수록 인식률이 좋아졌으며 제안된 방식의 4s (4s), 5 s의 쿼리에서는 평균 90% 이상의 좋은 인식결과가 나타나는 것을 볼 수 있다.

Table 3은 이퀄라이제이션(EQ), 잡음(NA), 피치-이동(PS), 타임-스트레칭(TS)의 환경에서의 실험 결과를 나타낸다. Table 3에서 나타난 것과 같이 방식 1은

162 이준용, 김형국

모든 환경에서 가장 좋은 인식률을 보여줬으며 방식 4는 피치-이동 시 다른 방식에 비해 상대적으로 좋은 결과를 보여준다.

Ⅳ. 결 론

본 논문은 MCLT 피크쌍 기반의 강인한 해시 함수를 이용하여 다양한 왜곡 환경에서의 강인한 오디오 핑거프린팅을 생성하는 방식에 대해 제안했다. 실험 결과 제안된 방식이 기존의 핑거프린트 알고리즘 보다 더 우수한 성능 가졌다는 것을 보여주며 이는 본 논문에서 제안한 핑거프린팅 방식이 잡음과 왜곡 환 경에서 강인하다는 것을 나타낸다.

또한 본 논문에서 제안된 방식은 저 연산의 장점을 가지고 있기 때문에 모바일 단말기에서 기존 방식보다 더 적합하게 사용 될 것이라 예상되며 특히스마트 티브이와 모바일 단말기 콘텐츠 검색 및 보안 애플리케이션에 적용될 것이라 사료된다.

향후 본 알고리즘을 최적화하는 방안과 비디오 핑 거프린트에 적용하는 방안에 대해 연구할 예정이다.

감사의 글

본 논문은 2013년도 정부(교육과학기술부)의 재원으로 한국연구재단 기초연구사업지원을 받아 수행된 것임(NRF-2013R1A1A2007601).

References

- W. Li, C. Xiao, and Y. Liu, "Low-order auditory Zernike moment: a novel approach for robust music identification in the compressed domain," EURASIP J. on Advances in Signal Processing, 1-15 (2013).
- A. Wang, "An industrial strength audio search algorithm," 4th International Conference Music Information Retrieval, 7-13 (2003).
- M. K. Mihcak and R. Venkatesan, "A perceptual audio hashing algorithm: A tool for robust audio identication and information hiding," In: Lett. Notes in Computer Science 2137, 51-65 (2001).
- T. Jiang, R. Wu, J. Li, K. Xiang, and F. Dai, "A real-time peak discovering method for audio fingerprinting," 5th International Conference on Internet Multimedia Computing

and Service, 368-371 (2013).

 S. Fenet, G. Richard, and Y. Grenier, "A scalable audio fingerprint method with robustness to pitch-shifting," 12th International Society for Music Information Retrieval Conference, 121-126 (2011).

▮ 저자 약력

▶이 준용 (Jun-Yong Lee)



2014년: 광운대학교 전자융합공학과 학사 2014년 ~ 현재: 광운대학교 전파공학과 석사과정

▶ 김 형 국 (Hyoung-Gook Kim)



1999년 ~ 2002년: 독일 SIEMENS/ Cortologic AG 책임연구원 2002년 ~ 2005년: 독일 베를린 공과대학교 Assistant Professor

2005년 ~ 2007년: 삼성종합기술원 수석 연구원

2007년~현재: 광운대학교 전자융합공학과 교수