

논문 2011-48SP-4-15

음향 데이터 전송 시스템의 강인한 데이터 검출 성능을 위한 Gaussian Mixture Model 기반 연구

(Data Detection Algorithm Based on GMM in the Acoustic Data
Transmission System)

송지현*, 장준혁**, 김문기***, 김동건***

(Ji-Hyun Song, Joon-Hyuk Chang, Moon-kee Kim, and Dongkeon Kim)

요약

본 논문에서는 패턴 인식에서 우수한 성능을 보여주는 가우시안 혼합 모델을 이용하여 MCLT 기반 음향 데이터 전송 시스템의 데이터 검출 성능 향상을 위한 방법을 제안하였다. 기존의 MCLT 기반 음향 데이터 전송 시스템에 대해서 분석하고, 이를 기반으로 데이터 검출 알고리즘에서 우수한 성능을 보여주는 특징 벡터를 선택하여 GMM의 입력 벡터로 효과적으로 이용한다. 다양한 음원(rock, pop, classic, jazz)과 마이크-스피커 사이의 거리 (1~5m)에서 시스템의 성능을 평가한 결과 GMM을 이용한 제안된 방법이 기존의 MCLT 기반 음향 데이터 전송 시스템의 데이터 검출 알고리즘보다 더욱 우수한 데이터 검출 성능을 보였다

Abstract

in this paper, we propose an approach to improve the data detection performance of the acoustic data transmission system based on the modulated complex lapped transform (MCLT). We first present an effective analysis of the features and the detection method of data in the acoustic data transmission system. And then feature vectors which are applied to the Gaussian mixture model (GMM) are selected from relevant parameters of the previous system for the efficient data detection. For the purpose of evaluating the performance of the proposed algorithm, Bit error rate (BER) of the received data was measured at different environments (music genres (rock, pop, classic, jazz) and different distances (1m~5m) from the loudspeaker to the microphone in a office room) and yields better results compared with the conventional scheme of the acoustic data transmission system based on the MCLT.

Keywords: Acoustic data transmission system, Modulated complex lapped transform (MCLT), Gaussian Mixture Model (GMM)

I. 서론

최근에는 추가의 모듈레이션이 필요 없이 단순히 휴대 단말기와 오디오 스피커만으로 은닉된 정보를 추출할 수 있는 음향 데이터를 이용한 음향 데이터 전송 시스템이 개발 되었다. 그 중 우수한 성능을 보여주는 음향 orthogonal frequency division multiplexing(OFDN)이 제안되었고, 음향 OFDN은 약 1600 Hz의 대역을 제거하고, 그 대역에 약 1kbps 정도의 데이터 전송이 가

* 학생회원 인하대학교 전자공학부
(Dep. of Electronics Engineering, Inha University)

** 정회원, 한양대학교 융합전자공학부
(Dep. of Electronics Engineering, Inha University)

*** 정회원, C&I Biz 미래기술원 AT개발팀, SK 텔레콤
(C&I Biz Future Tech. Institute, SK Telecom)

※ 이 논문은 2009년 정부(교육과학기술부)의 재원으로
한국연구재단의 지원을 받아 수행된 연구임
(2009-0085162) 또한 SK텔레콤의 지원을 받아 수행
된 연구임

접수일자: 2011년1월20일, 수정완료일: 2011년2월9일

능한 OFDM 신호를 삽입하여 은닉된 데이터를 전송한다^[1, 3]. 하지만 음향 OFDM의 경우 정보를 삽입하기 위해서 일부 대역을 필터로 걸러낸 후 그 대역에 정보를 삽입하고, 프레임 간 중첩이 되지 않기 때문에 필연적으로 음질 저하 문제가 발생한다. 최근에는 이러한 음질 저하 문제를 개선한 modulated complex lapped transform (MCLT) 기반의 음향 데이터 전송 시스템이 제안되었다^[4~5]. 제안된 기술은 정보를 삽입하기 위해서 오디오 신호를 MCLT 변환한 후 MCLT 계수의 위상을 변화 시켜서 데이터를 삽입한다. 여기서 MCLT의 중첩에 의해서 데이터의 위상이 급격히 변하는 것을 방지하여 음질이 저하되는 현상을 줄였다^[6, 8]. 하지만, 실제 실험 환경 즉, 휴대폰과 오디오 스피커 사이의 거리, 방 사이즈, 장애물 존재 유무 등과 같은 외부 요인에 의해서 MCLT의 위상이 영향 받기 때문에 변화된 실험 환경에서 보장된 검출 성능을 보여주기 위해서 새로운 검출 알고리즘이 필요하다.

본 논문에서는 MCLT 기반의 음향 데이터 전송 시스템의 검출 성능을 향상시키기 위하여 Gaussian mixture model (GMM) 기반의 데이터 검출 알고리즘을 제안한다. 기존의 시스템에서 데이터 검출을 위해 사용된 특징벡터 외에 추가로 통계적 분류 특성이 우수한 특징벡터를 선택하여 GMM에 적용한 결과 제안된 GMM기반의 데이터 검출 성능이 우수한 성능을 보임을 알 수 있었다.

본 논문의 구성으로는, II장에서는 MCLT 기반 음향 데이터 전송 시스템의 개요 및 데이터 삽입, 추출 방법에 대해서 기술하였고, III장에서는 제안된 GMM 기반의 데이터 검출 알고리즘에 대해서 기술한다. IV 장에서는 실험 결과에 대한 비교 및 분석에 대해 기술하였으며, 마지막으로 V장에서 결론을 맺는다.

II. MCLT 기반의 음향 데이터 전송 시스템

MCLT 기반의 음향 데이터 전송 시스템은 오디오 신호에 인지할 수 없는 신호를 삽입하고, 음향 신호를 통해서 정보를 전달하는 시스템으로, 시간 축 상의 오디오 신호를 주파수 축으로 바꾸기 위해서 MCLT를 사용한다^[3~4]. MCLT는 2M 길이의 시간 축 신호를 입력으로 받아 M 길이의 주파수 신호로 바꾸어 주고, 역변환시 인접한 MCLT 프레임간의 중첩을 통해서 오디오 신호가 구해진다. 그림 1은 MCLT 기반의 음향 데이터 전송 시스템의 블록 다이어그램을 나타내고, 인코딩 과정은 다음과 같다. 먼저 입력된 오디오 신호를 MCLT를 통해서 주파수 신호로 변환하게 된다. 변환된 MCLT 계수는 복소수 값이기 때문에 실수 성분과 허수 성분을 가지고 있어 위상 값을 얻어 낼 수 있다. 보내고자 하는 데이터를 binary phase shift key (BPSK) 방식을 통해서 비트열로 바꾸고, 선택된 MCLT 계수의 위상 변조를 통해서 삽입한다. 수정된 주파수 신호는 IMCLT를 통해서 오디오 신호로 바꾸고, 인접한 MCLT 프레임끼리 중첩하여 더해준다. 정보가 삽입된 신호는 스피커를 통해서 전송되고, 휴대폰과 같은 마이크로폰을 통해서 입력받는다. 입력된 신호는 동기화 과정을 통해서 데이터의 위치를 찾고, MCLT 변환을 통해서 다시 주파수 신호로 변환시킨다. 다음으로 정보가 삽입된 주파수 계수의 위상을 통해서 숨겨진 데이터를 검출한다.

1. 데이터 삽입 방법

사운드 코드 시스템에서 데이터를 오디오 신호에 삽입하기 위해서 데이터를 BPSK로 변조한다. 그 후 그림 2에 나타내어진 것처럼 선택되어진 MCLT 계수의 위상

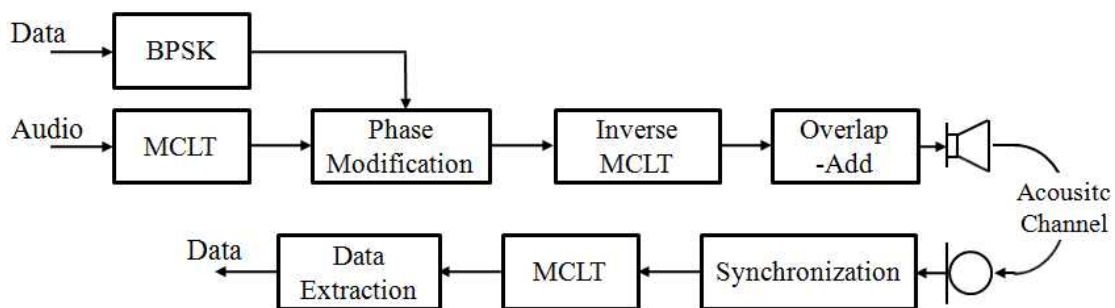


그림 1. MCLT 음향 데이터 전송 시스템 블록 다이어그램

Fig. 1. Block diagram of the acoustic data transmission system based on the MCLT.

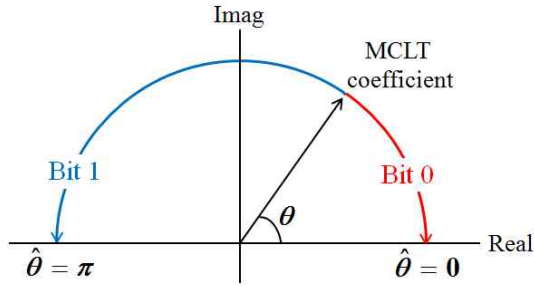


그림 2. MCLT 음향 데이터 전송 시스템 데이터 삽입 방법

Fig. 2. Data Embedding.

을 삽입 정보에 따라서 0 또는 π 로 바꾸어 준다.

$$\hat{X}_i(k) = \begin{cases} |X_i(k)|d_i(k) & \text{if } k \in D \\ X_i(k) & \text{otherwise} \end{cases} \quad (1)$$

식(1)은 데이터 정보 삽입 방식을 수식으로 나타낸 것이다. 여기서 $\hat{X}_i(k)$ 는 정보가 삽입된 MCLT 계수를 나타내고, i 는 프레임 인덱스, k 는 주파수 인덱스, D 는 정보가 삽입될 주파수 인덱스의 집합 (6.4 kHz ~ 8 kHz, 총 36개의 주파수 밴드)을 나타낸다. $d_i(k)$ 는 k 번째 부반송파에 삽입하는 BPSK로 변조된 데이터 비트로 -1 (bit 0) 또는 1 (bit 1)의 값을 갖는다.

사운드 코드의 검출 성능을 높이기 위해서 대역 확산을 적용하였다^[9]. 길이 L 인 확산 부포를 \mathbf{s} 라고 하면 대역 확산된 데이터는 다음과 같이 삽입될 수 있다.

$$\begin{aligned} \hat{X}_i(mM+k) &= \begin{cases} |X_i(mM+k)|d_i(k)s(m) & \text{if } mM+k \in D \\ X_i(k) & \text{otherwise} \end{cases} \end{aligned} \quad (2)$$

여기서 $m = 0, 1, 2, \dots, L-1$ 의 값을 갖고, M 은 한 개의 프레임에 삽입되는 비트열의 길이를 나타낸다. 대역 확산을 사용할 경우 비트 전송률이 대역확산 길이(L)에 반비례 하지만 검출 성능이 더욱 우수하기 때문에 비트 전송률과 검출 성능과의 관계를 고려하여 L 값을 조정하여 준다. 실제 음향 데이터 전송 시스템에서는 $M = 9$, $L = 4$ 로 사용하여 한 프레임에 $M * L = 36$ bit의 확산된 데이터가 삽입된다.

2. 데이터 검출

데이터의 검출은 동기 신호가 정확하게 검출되었을 때 다음과 같은 판별식을 통해서 구해진다.

$$\sum_{m=0}^{L-1} \text{Re}\{\hat{Y}_i(mM+k)\}s(m) \begin{cases} > 0 & \text{bit 0} \\ < 0 & \text{bit 1} \end{cases} \quad (3)$$

여기서 $\hat{Y}_i(k)$ 는 수신된 i 번째 데이터 프레임의 MCLT 계수를 나타낸다. 대역 확산 부호와 수신된 MCLT 계수 $\hat{Y}_i(k)$ 의 실수 값을 이용하여 구해진 결과 값이 0보다 크면 bit 0을 0보다 작으면 bit 1로 판단한다. 즉, 데이터 검출은 $\hat{Y}_i(k)$ 의 실수 값만을 이용하여 구해진다.

III. 제안된 GMM 기반의 검출 시스템

음향 신호의 경우 실험 환경 (방 사이즈, 장애물 존재 유무, 오디오 스피커와 마이크 사이의 거리)과 같은 외부 요인에 의해 MCLT의 위상이 영향 받는다. 따라서 식 (3)과 같이 고정된 파라미터를 이용한 검출 방식은 음향 데이터 전송 시스템의 성능 저하를 발생시킨다. 그림 3과 4는 실제 실험을 통해서 수신된 MCLT 계수의 실수 값과 허수 값에 대한 히스토그램을 나타낸다. 그림 3을 통해서 식 (3)에 의해 비트 정보를 검출하였을 경우 즉, 수신 MCLT의 실수 값을 기준으로 비트 정보를 판단하였을 경우 실수 값의 0을 기준으로 오른쪽 실선과 왼쪽 점선에서 비트 오류가 발생하는 것을 쉽게 확인할 수 있다. 다음으로 그림 4를 통해서 수신 MCLT 계수의 허수 값 또한 bit 0과 bit 1을 분류하는

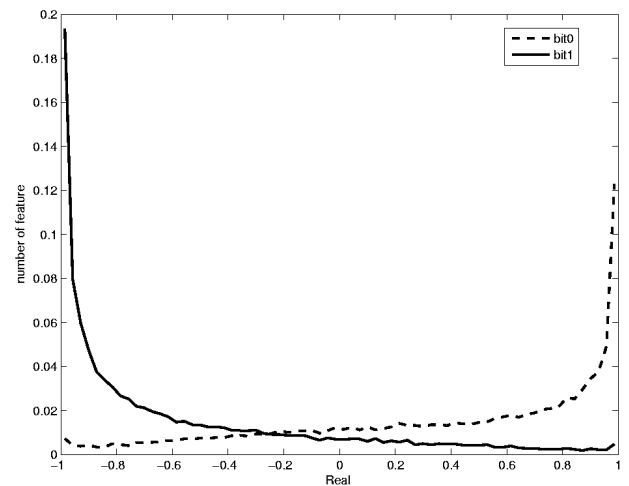


그림 3. MCLT 계수 실수 값의 히스토그램

Fig. 3. Histogram of the real part of the MCLT coefficient.

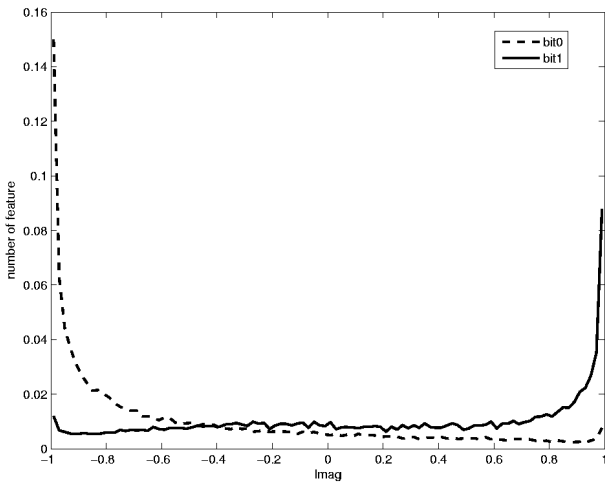


그림 4. MCLT 계수 허수 값의 히스토그램
Fig. 4. Histogram of the imaginary part of the MCLT coefficient.

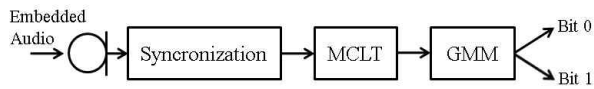


그림 5. 제안된 데이터 검출 알고리즘 블록 다이어그램
Fig. 5. Block diagram of the proposed method.

데 우수한 특징 벡터가 될 수 있음을 알 수 있다. 따라서 음향 데이터 전송 시스템의 검출 성능을 향상시키기 위해서 수신 MCLT 계수의 실수 값뿐만 아니라 허수 값 또한 이용하는 분류 알고리즘을 제안한다. 특히, 특징 벡터의 통계적 편차가 클수록 더욱 우수한 성능을 보여주는 GMM (Gaussian mixture model)을 적용하였고, 그림 5에 제안된 알고리즘의 블록다이어그램을 나타내었다.

그림 3을 통해서 식 (3)에 의해 비트 정보를 검출하였을 경우 즉, 수신 MCLT의 실수 값만을 기준으로 비트 정보를 판단하였을 경우 실수 값의 0을 기준으로 오른쪽 실선과 왼쪽 점선에서 비트 오류가 발생하는 것을 쉽게 확인할 수 있다. 다음으로 그림 4를 통해서 수신 MCLT 계수의 허수 값 또한 bit 0과 bit 1을 분류하는데 우수한 특징 벡터가 될 수 있음을 알 수 있다. 따라서 음향 데이터 전송 시스템의 검출 성능을 향상시키기 위해서 수신 MCLT 계수의 실수 값뿐만 아니라 허수 값 또한 이용하는 분류 알고리즘을 제안한다. 특히, 특징 벡터의 통계적 편차가 클수록 더욱 우수한 성능을 보여주는 GMM (Gaussian mixture model)을 적용하였고, 그림 5에 제안된 알고리즘의 블록다이어그램을 나타내었다.

GMM은 가우시안 혼합성분 밀도의 가중치 합인 함수로서 다음과 같이 정의 된다^[10-11].

$$P(x|\lambda) = \sum_{i=1}^M \alpha_i P_i(x) \tag{4}$$

여기서

$$P_i(x) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma_i|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(x-\mu_i)^T (\Sigma_i)^{-1} (x-\mu_i)\right\} \tag{5}$$

여기서 α_i 는 혼합 성분의 가중치를 나타내고 μ_i 는 평균 벡터, Σ_i 는 공분산 행렬을 나타낸다.

$$\lambda = \{\alpha_i, \mu_i, \Sigma_i\}, \quad i = 1, \dots, M. \tag{6}$$

GMM은 크게 bit 0과 bit 1에 대한 특징 벡터의 분포를 가장 잘 나타낼 수 있는 모델을 찾는 훈련부와 이 모델을 이용하여 인식하는 인식부로 나눌 수 있다. 먼저 훈련 부는 위와 같은 파라미터를 가지고 Expectation Maximization (EM) 알고리즘 기반의 학습을 통해서 bit 0/ bit1에 대한 혼합 가우시안 모델 λ 을 추정하고 인식부는 훈련부에 만들어진 λ 을 이용해서 수신된 MCLT 계수에 대한 사후 확률을 구하여 가장 큰 확률을 갖는 모델을 찾는다.

$$\hat{S} = \arg \max_{k \in \{bit0, bit1\}} \sum_{t=1}^T \log p(\vec{x}_t | \lambda_k) \tag{7}$$

본 실험에서는 GMM의 특징 벡터로서 대역 확산된 MCLT 계수의 실수 값과 허수 값을 특징벡터로 사용하여 다양한 혼합성분 개수로 λ_{bit0} (bit 0 모델)과 λ_{bit1} (bit 1 모델)를 추정하였다. 테스트 과정에서 GMM의 모델 파라미터 λ_{bit0} 과 λ_{bit1} 에 실제 데이터가 입력되어 각 모델에 대한 우도를 생성하고, 아래와 같은 우도비 (Likelihood Ratio)를 이용하여 수신 데이터를 검출한다.

$$A_x^{(t)} = \frac{P(\vec{x}^{(t)} | \lambda_{bit0})}{P(\vec{x}^{(t)} | \lambda_{bit1})} \begin{matrix} > \\ < \end{matrix} \begin{matrix} bit0 \\ bit1 \end{matrix} \eta \tag{8}$$

여기서 η 는 bit 0/ bit 1 분류의 문턱 값이고, t 는 프레임 번호를 나타낸다.

IV. 실험

본 논문에서는 제안된 GMM 기반의 데이터 검출 성능을 평가하기 위해서 BER 기반의 검출 성능 평가에 대한 실험을 하였다. 실험에 사용된 음향 데이터 전송 시스템의 실험 환경 파라미터는 표 1과 같다.

GMM 훈련을 위해서 1~5 m에서 녹음된 rock, pop, classic, jazz 음악 파일이 사용되었고, 신뢰성 있는 결과 도출을 위해서 훈련에 사용된 데이터는 테스트에 사용되지 않았다. 주파수 대역에 따라서 위상 변화 결과가 틀리게 나오기 때문에 각 주파수 대역에 대해서 각각 모델 파라미터를 추출하였다. 즉, GMM의 특징 벡터로 대역 확산된 4개의 MCLT 계수의 실수 값과 허수 값을 이용하여 총 8개로 구성하였고, 그림 6에 나타내었다. 비트 검출을 위한 주파수 대역은 9개 (36(삽입 주파수 대역)/4(대역 확산 길이))를 사용하였다. 따라

표 1. 음향 데이터 전송 시스템 실험 환경 파라미터
Table 1. The acoustic data transmission system parameters.

샘플링 주파수	44.1 kHz
데이터 삽입 주파수 대역	6.4 ~ 8 kHz
MCLT 프레임 크기	512 샘플
데이터 블록 길이	40 프레임
삽입 주파수 대역 수	36
대역확산 길이	4

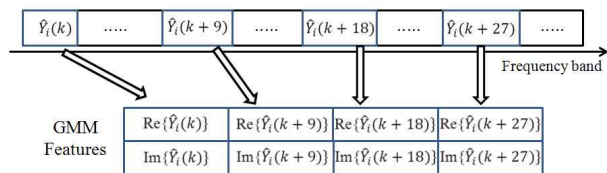


그림 6. GMM 특징벡터
Fig. 6. Feature vectors of the GMM.

표 2. 비트오류 테스트
Table 2. BER of the data transmission systems.

Method	Distance	Rock	Pop	Classic	Jazz	Average
Proposed	1m	0.029	0.041	0.115	0.107	0.073
	2m	0.079	0.051	0.162	0.110	0.100
	3m	0.095	0.124	0.241	0.198	0.164
	4m	0.136	0.111	0.288	0.238	0.193
	5m	0.154	0.158	0.269	0.199	0.195
Paper [4]	1m	0.086	0.095	0.121	0.117	0.105
	2m	0.090	0.111	0.179	0.156	0.134
	3m	0.101	0.132	0.243	0.186	0.166
	4m	0.152	0.183	0.265	0.213	0.203
	5m	0.119	0.146	0.247	0.203	0.179

서 총 모델 파라미터는 16개 (bit 0 8개, bit 1 8개)가 된다. GMM의 혼합 성분 개수는 검출 성능과 계산량을 고려하여 16차로 사용하였다.

데이터 검출 성능을 비교 평가하기 위해서 오디오 스피커와 마이크 사이의 거리 1~5 m에서 rock, pop, classic, jazz 에 대해서 녹음하였고, 비트 오류율을 측정 하였다. 비트 오류율은 동기화 신호가 정확하게 검출 되었을 때 추출된 데이터 삽입 대역의 MCLT 계수를 이용하여 구해진 결과와 매뉴얼의 비교를 통해서 구해졌고, 표 2는 측정 결과를 나타낸다.

BER 테스트 결과 기존 음향 데이터 전송 시스템의 성능보다 제안된 GMM 기반의 검출 알고리즘이 더욱 우수한 성능을 보여주는 것을 볼 수 있다. 특히, rock, pop의 오디오 파일의 경우 거리 1m에서 평균 5.7%, 5.4% 정도의 성능 향상을 보여주었다. 실험 결과를 통해서 확인할 수 있듯이 수신된 MCLT 계수의 실수 값뿐만 아니라 허수 값 또한 비트 검출을 위한 중요한 특징 벡터가 될 수 있음을 알 수 있다.

VI. 결론

본 논문에서는 사운드코드의 비트 검출 성능을 향상시키기 위해 GMM 기반의 비트 검출 방법을 제안하였다. GMM의 특징 벡터로 대역 확산된 4개의 MCLT 계수 실수 값과 허수 값을 적용하였고, 기존의 사운드코드와 비트 검출 성능을 비교한 결과 제안된 GMM 기반의 비트 검출 알고리즘이 더욱 우수한 성능을 보여주는 것을 알 수 있었다. 추후 적용적 GMM을 이용하여 GMM의 모델 파라미터를 업데이트 해준다면 거리에 따른 비트 검출 율을 더욱 향상시킬 수 있을 것이다.

참고 문헌

- [1] H. Matsuoka, Y. Nakasuma and T.Yoshimura, "Acoustic Communication system using mobile terminal microphones," *NTT DoCoMo Technica journal*, Vol. 8, no 2, pp. 2-12, Sep. 2006.
- [2] H. Matsuoka, Y. Nakasuma and T.Yoshimura, "Acoustic Communication With OFDM signal Embedded in Audio", *AES 29th International Conference*, pp.1-6, Sep. 2006.
- [3] K. Cho, H. S. Yun, J. -H. Chang and N. S. Kim, "An analysis on audio quality deterioration of acoustic OFDM," *the journal of the Acousical*

Society of Korea, Vol. 28. no. 2, pp. 107-111. Feb. 2009.

[4] K. Cho, H. S. Yun, S. S. Kim and N. S. Kim, "Imperceptible data hiding in MCLT domain for acoustic data transmission using loudspeaker and microphone", *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, pp. 238-241, Oct. 2009.

[5] H. S. Yun, K. Cho and N. S. Kim, "Acoustic data transmission based on modulated complex lapped transform," *IEEE Signal Processing Lett.*, Vol. 17, no. 1, pp. 67-70, Jan. 2010.

[6] H. S. Malvar, "Modulated Complex Lapped Transform for Integrated Signal Enhancement and Coding," *U.S. Patent 6 496 795*, Dec. 2006.

[7] D. Qionghai and C. Xinjian, "New algorithm for modulated complex lapped transform with symmetrical window function," *IEEE Signal Process. Lett.*, Vol. 11, no. 12, pp. 925-928, Dec. 2004.

[8] H. S. Malvar, "Fast algorithm for the modulated complex lapped transform," *IEEE Signal Process. Lett.*, Vol. 10, no. 1, pp. 8-10, Jan. 2003.

[9] D. Kirovski and H. S. Malvar, "Spread-Spectrum Watermarking of AudioSignals", *IEEE Trans. on Signal Proc.*, Vol. 51, no.4, pp.1020-1033, Apr. 2003.

[10] G. Xuan, W. Zhang and P. Chai, "EM algorithm of gaussian mixture model and hidden Markov model," *Proc. IEEE International Conference on Image Processing*, vol. 1, pp. 145-148, Oct. 2001.

[11] D. A. Reynolds, T. F. Quatieri, and R. B. Dunn, "Speaker verification using adapted Gaussian mixture models," *Digital Signal Processing*, vol. 10, pp. 19-41, Jan. 2000.

— 저 자 소 개 —



송 지 현(학생회원)
 2007년 인하대학교 전자전기공학부 학사.
 2009년 인하대학교 전자공학과 석사
 2010년 인하대학교 전자공학과 박사과정.

<주관심분야 : 음성신호처리, 디지털신호처리, 잡음제거>



김 문 기(정회원)
 1989년 경북대학교 전기전자공학과 학사
 1989년~1994년 KT 정보통신사업본부
 1994년~현재 SK텔레콤

<주관심분야 : 무선망 엔지니어링, 방송통신 융합 기술, 휴먼인터페이스 기술>



장 준 혁(정회원)
 1998년 경북대학교 전자공학과 학사.
 2000년 서울대학교 전기공학부 석사.
 2004년 서울대학교 전기컴퓨터공학부 박사.

2000년~2005년 (주)벤티스 연구소장
 2004년~2005년 캘리포니아 주립대학, 산타바바라(UCSB) 박사후연구원
 2005년 한국과학기술연구원(KIST) 연구원
 2005년~2011년 인하대학교 전자공학부 조교수
 2011년~현재 한양대학교 융합전자공학부 부교수
 <주관심분야 : 음성신호처리, 오디오 신호처리, 통신 신호처리, 휴먼/컴퓨터 인터페이스>



김 동 건(정회원)
 2007년 서울대학교 기계항공공학부 학사
 2009년 서울대학교 기계항공공학부 석사
 2009년~현재 SK텔레콤

<주관심분야 : 방송통신 융합 기술, 음향 통신 기술>