

비음수 제약을 통한 일반 소리 분류

(Classification of General Sound with Non-negativity Constraints)

조 옹 춘 [†] 최 승 진 ^{**} 방 승 양 ^{***}

(Yong-Choon Cho) (Seungjin Choi) (Sung-Yang Bang)

요 약 전체관적인 표현방법인 최소 코딩 또는 독립 성분 분해(ICA)는 이전의 청각의 처리와 소리 분류의 작업을 해명하는데 성공적으로 적용되었다. 반대로 부분 기반 표현법은 뇌에서 물체를 인식하는 방법을 이해하는 또 다른 방법이다. 이 논문에서, 우리는 소리 분류의 작업에 부분기반 표현법을 학습시키는 비음수화 행렬 분해(NMF)[1] 방법을 적용하였다. 잡음이 존재할 때와 존재하지 않을 때 두 가지 상황에서, NMF를 이용하여 주파수-시간영역의 소리로부터 특징을 추출하는 방법을 설명한다. 실험결과에서는 NMF에 기반을 둔 특징이 ICA에 기반을 두어 추출한 특징보다 소리 분류의 성능을 향상시킴을 보여준다.

키워드 : 소리분류, 비음수화 행렬 분해, 오디오 신호 처리

Abstract Sparse coding or independent component analysis (ICA) which is a holistic representation, was successfully applied to elucidate early auditory processing and to the task of sound classification. In contrast, parts-based representation is an alternative way of understanding object recognition in brain. In this thesis we employ the non-negative matrix factorization (NMF) which learns parts-based representation in the task of sound classification. Methods of feature extraction from the spectro-temporal sounds using the NMF in the absence or presence of noise, are explained. Experimental results show that NMF-based features improve the performance of sound classification over ICA-based features.

Key words : Sound Classification, Non-negative Matrix Factorization, Audio signal processing

1. 서 론

소리 분류는 많은 흥미로운 적용을 가지는 오디오 신호처리에서 매우 중요한 문제이다. 예를 들어 음성/비음성 분류는 자동 음성 인식 시스템의 성능 향상에 사용될 수 있다. 또한 오디오 신호를 음성, 음악, 배경 소리 등의 다양한 형태로 구분하는 것은 오디오 검색 시스템에 유용하다. 대부분의 소리분류 시스템들은 주파수나 스펙트럼에 기반을 둔 특징을 사용한다. 하지만, 직접적인 스펙트럼 기반 특징은 그 자체의 높은 차원과 지각적으로 비슷한 신호들의 현저한 분산[2]으로 인해서 소리 분류에서는 적합하지 않다. 최근에 Casey는 MPEG-7에 채택된 ICA 기반 소리 분류 시스템을 제안하였다[2,3].

ICA는 다변수 데이터를 통계적으로 독립인 계수를 갖는 비 직교 기저벡터들의 선형적인 조합으로 분해하는 통계적인 방법이다[4,5]. ICA는 효율적인 부호화의 관점에서 초기의 청각의 처리를 해명하는데 적용되었으며[6,7], 최소 청각 수용 체를 설명하는데 잘 맞는 것으로 보였다[8]. 또한 ICA는 감각 정보를 효과적으로 부호화 하는 방법이고, 최소 코딩의 방법으로 그 효율성은 초기의 청각 및 시각 시스템을 설명하는데 증명되었다[9-11]. 비록 ICA가 자연적인 소리의 더 높은 차원의 통계적 구조를 학습하지만, 이는 기준 벡터들이 양 또는 음의 계수들의 조합을 허용하므로 전체관적인 표현법이라 할 수 있다.

부분 기반 표현법은 뇌에서의 인식을 이해하는 또 다른 방법이고, 어떤 계산 이론들은 그러한 표현방법에 의존하고 있다. 예를 들어서, Biederman은 어떤 물체라도 *geons*(geometric ions)이라고 언급되는 지각적 알파벳의 개념으로 설명될 수 있다고 주장한다[12]. 부분 기반 표현을 학습시키기 위한 직관적인 아이디어는 기저 벡터의 선형 결합에 감산이 없도록 제한하는 것이다.

[†] 비 회 원 : 삼성전자 DM총괄 DM연구소 Audio Lab
yongc@postech.ac.kr

^{**} 정 회 원 : 포항공과대학교 컴퓨터공학과 교수
seungjin@postech.ac.kr

^{***} 중신회원 : 포항공과대학교 컴퓨터공학과 교수
sybang@postech.ac.kr

논문접수 : 2003년 12월 23일

심사완료 : 2004년 7월 10일

NMF[13]는 지각 데이터의 부분기반 표현에 대한 간단한 승산식 갱신 알고리즘이다.

이 논문에서 우리는 NMF를 이용한 소리 분류 방법을 제안한다. 제안하는 소리 분류 시스템은 특징으로 주 파수-시간 영역의 소리들로부터 비음수 요소 부분들을 추출해 낸다. NMF를 이용해 계산해낸 기저벡터들은 다시 정렬되고, 그것들 중에서 일부분이 그 자체의 분별력에 따라서 선택된다. 추출된 소리 특징들은 이런 축소된 벡터들로부터 계산되고, 이는 은닉 마르코프 모델(HMM) 분류기의 입력으로 들어간다. 또한, 부가적인 잡음에 강건한 소리의 특징을 학습하는 간단한 방법을 소개한다. 마지막으로 제안한 방법과 ICA에 기반을 둔 방법과 비교하고 제안한 방법이 성능 면에서 우수함을 보여준다.

2. 비음수화 행렬 분해

효과적인 정보 표현법은 감각 데이터의 인식을 이해하고 패턴분류에 있어서 중요한 역할을 한다. 초기의 청각 신호 처리 분야에서 효과적인 코딩 전략을 설명하는 한 가지 방법은 외부 세계에서 나온 신호들의 구조는 기저 함수들의 선형 중첩의 개념으로 모델 되는 선형 발생 모델에 기반을 둔다. 다시 말해서 선형 발생 모델은 관찰된 데이터 $x_t \in m$ 는 아래와 같은 수식으로 표현된다.

$$x_t = A s_t + \epsilon_t \tag{1}$$

여기서 $A \in m \times n$ 는 열벡터로 기저벡터 $a_i \in m$ 를 포함하며, $s_t \in n$ 는 잠재변수이며, $\epsilon_t \in m$ 은 데이터 모델에서 불확정성을 나타내는 잡음 벡터이다. 선형 발생 모델 학습을 위한 여러 가지 방법들은 인자 분해를 포함하여, 주 성분 분석(PCA), 희소 코딩, 독립 성분 분석(ICA) 등이 있다. 일반적으로 이러한 방법들은 전체과적인 표현법이라 할 수 있다.

반면에 뇌에서 부분 기반 표현법을 사용한다는 몇몇의 증거가 있으며, 물체 인식의 어떤 계산 이론에서는 이러한 표현법들을 이용한다. 선형 발생 모델 (1)을 이용하여 부분 기반 표현법을 찾아내는 한 가지 방법은 관찰된 데이터를 형상화 하는데 사용되는 기저 벡터들의 비 차감 조합이 없도록 기저 벡터들과 잠재변수들 전부에 비음수 제약을 가하는 것이다. 비음수화 행렬 분해(NMF)[1]는 비음수 제약을 통해 선형 데이터 표현법을 찾아내는 부분공간 방법이다.

N 개의 관찰된 데이터 포인트들 $\{x_t\}, t=1, \dots, N$ 가 있다고 가정하고, 데이터 행렬을 $X = [x_1 \dots x_N]$ 이라 하자. 잠재변수 행렬 S 도 비슷한 방법으로 정의한다. 푸아송 잡음 모델에 의해 로그-유사도는 아래와 같다.

$$L = \sum_{i=1}^N \sum_{j=1}^m \{ X \log(A S) - (A S) \} \tag{2}$$

(2)의 국지적 최대 값은 아래와 같은 승산식 갱신 규칙에 의해서 찾을 수 있다[13].

$$S_{a\mu} \leftarrow S_{a\mu} \frac{\sum_i A_{ia} X_{i\mu} / (A S)_{i\mu}}{\sum_k A_{ka}} \tag{3}$$

$$A_{ia} \leftarrow A_{ia} \frac{\sum_i A_{ia} X_{i\mu} / (A S)_{i\mu}}{\sum_j S_{aj}} \tag{4}$$

A 와 S 의 모든 항목은 비음수이며, 따라서 비 차감 조합만이 허용된다. 이는 전체로부터의 부분들의 조합의 직관적인 개념으로 호환되는 것으로 믿어지며, NMF가 부분기반 표현법을 학습하는 방법[1]이다. 또한 점화율은 비음수라는 생물학적인 사실과도 연관되어진다.

3. NMF를 통한 소리의 특징 추출

오디오 신호에서 특징을 추출하는 방법은 세 가지 단계로 구성된다. 첫째로, 스펙트로그램을 계산하고 시간 축을 따라서 이미지 조각의 연속으로 잘라낸다. 이는 데이터 행렬 X 가 되며, NMF에 의해 기저 행렬 A 와 부호화 행렬 S 의 곱으로 분해 된다. 다음으로 기저 벡터들의 분별력에 따라 몇몇의 기저 벡터가 선택되고, 마지막으로 이렇게 선택된 기저 벡터들을 이용하여 특징들이 학습된다. 전체적인 구조적 다이어그램은 그림 1과 같다.

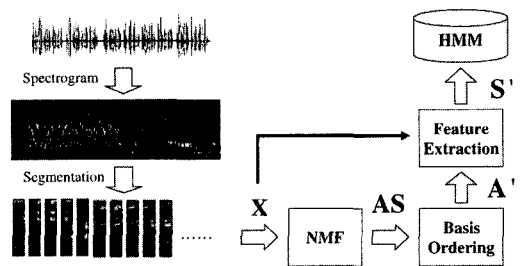


그림 1 제안한 소리 분류 시스템 전체 구조

3.1 스펙트럼 신호의 비음수화 분해

시간 영역에서 표본된 오디오 신호들은 시간과 관련된 스펙트럴 에너지를 표현하는 스펙트로그램으로 변환된다. 스펙트로그램은 시간 축을 따라서 이미지 조각들로 나뉜다. 그러므로 시간 영역의 오디오 신호 대신에 음수 값을 허용치 않는 이미지 순차의 집합을 가지고 분석하게 된다. 각 이미지 조각들은 벡터로 변환해서 N 개의 m 차원 벡터로 구성된 데이터 행렬 X 를 만든다.

NMF는 $X \in \mathbb{R}^{m \times N}$ 를 기저 행렬 $A \in \mathbb{R}^{m \times n}$ 와 부호화 행렬 $S \in \mathbb{R}^{n \times N}$ 의 곱으로 분해한다. 부호화 변수(기저 계수)들의 개수 n 은 관찰된 데이터의 수 m 보다 작게 설정한다. 다시 말하면, 스펙트로그램의 각 이미지 조각은 관련된 기저 이미지들에 대한 공헌도를 나타내는 부호화 변수들과 국소화된 기저 이미지의 선형 중첩으로 나타낼 수 있다. 그림 2는 예로 NMF와 ICA를 통해 계산된 기저 이미지들을 보여준다. NMF 기저 이미지는 ICA를 통한 기저 이미지보다 훨씬 더 국지적인 특징이 존재한다. NMF와 ICA는 모두 원칙적으로 희소 코딩과 관련되어 있지만, NMF에 의한 부분 기반 표현은 비음수 데이터에 대해 훨씬 더 국지적이고 희소 특징을 이끌어 낸다.

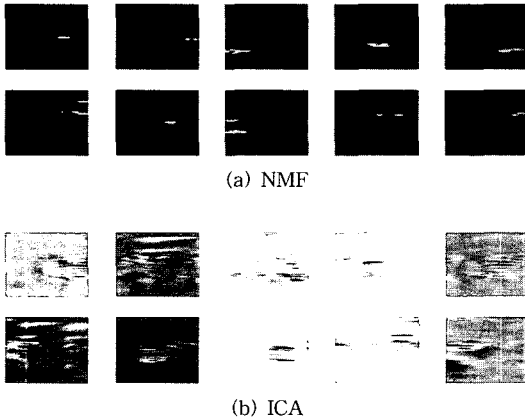


그림 2 스펙트로그램에서 NMF와 ICA를 통해 학습된 10개의 기저 이미지(총 150개)

3.2 기저 선택

NMF는 기저 이미지가 부류 라벨에 관계없이 학습되는 비 관리자 학습방법이다. 하지만 부류 정보는 학습 단계에서는 가능하기 때문에 이 정보를 고려하는 것이 타당하다. 제안하는 기저 선택법은 아래와 같이 정의하는 분별력에 기반을 둔다.

$$J(k) = \sum_i \sum_j \frac{|m_{ik} - m_{jk}|}{\sigma_{ik} + \sigma_{jk}}, \quad 1 \leq k \leq n \quad (5)$$

여기서 m_{ik} 와 σ_{ik} 는 부류 i 에 관련되는 행렬 S 의 k 번째 행벡터의 평균과 분산을 나타낸다. 분별력 (5)의 수식은 차이가 많은 평균과 작은 분산을 요구하는 FLD (Fisher's Linear Discriminant)의 방법을 연상시킨다. 그림 3은 150개의 기저 벡터에 대한 분별력 측정치를 보여준다. 여기에서 x -축은 기저 벡터의 번호를 나타내고, y -축은 측정된 분별력의 크기를 나타낸다. 적절한 경계 값을 설정하여 더 좋은 분별력을 가질 수 있을 만

큼의 기저 벡터를 골라낸다. 이러한 분별력 측정 도구를 이용하여 $x \leq n$ 개의 기저 벡터를 갖는 행렬 A' 를 선택한다.

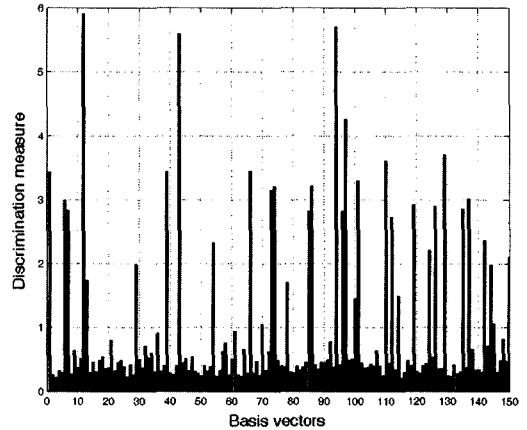


그림 3 각 기저 벡터에 대한 분별력 측정

3.3 특징 학습

NMF 기저 이미지들은 주파수 영역뿐만 아니라 시간 영역에서도 국지화 되어있는 청각 수용 체의 특징을 보여준다(그림 2 참조). 비록 NMF가 선형 데이터 모델을 사용하지만, 기저 행렬 A 와 관찰된 데이터 x 가 주어졌을 때 은닉 변수 s 의 추정 은 비음수 제약 때문에 비선형 절차를 따르게 된다. 따라서 A 와 x 가 주어졌을 때 최적의 은닉 변수를 찾아내는 것은 쉬운 일이 아니다.

이 논문에서 NMF에 의해 계산된 n 개의 기저 벡터에서 분별력 측정법 (5)를 이용하여 선택된 k 개의 기저 벡터로 구성된 $A \in \mathbb{R}^{m \times n}$ 이 주어졌을 때, 최적의 은닉 변수를 추정해내기 위해서 두 가지 방법을 살펴본다.

방법 I: A' 와 관련된 부호화 변수 행렬 S' 을 계산하기 위해서, A' 을 고정시키고 갱신 규칙 (3)을 이용한다.

방법 II: 방법 I에서 오직 선택된 기저 벡터들만 수식 (3)에 의해서 관련된 부호화 변수를 추정하는데 사용된다. 다시 말하자면, NMF에 의해 계산된 $n-x$ 의 기저 벡터들은 부호화 변수를 추정하는데 어떤 기여도 하지 않았다는 것이다. 반면에 방법 II에서는 부호화 변수 행렬 S' 을 추정하는데 $n-x$ 개의 기저 벡터들 A'' 을 사용한다. 기저 행렬 A 는 $A = [A', A'']$ 으로 놓을 수 있고, A' 은 고정시키고 A'' 만 갱신 식에 의해서 갱신된다. 그러면, 부분적으로 갱신되는 행렬 A 는 새로운 부호화 변수 행렬 S 를 추정하는데 사용된다. 결과적으로 A' 과 연관되는 $S = [S'^T, S''^T]$ 의 일부인 S' 만

이 분류를 위해 사용된다. 이 과정을 요약하면 아래의 수식과 같다.

$$A_{ia}^{\leftarrow} = A_{ia}^{\rightarrow} \frac{\sum_{\mu} S_{a\mu} X_{i\mu} / (A^{new} S)_{i\mu}}{\sum_{\nu} S_{a\nu}} \quad (6)$$

$$A^{new} = [A', A''], \quad A'' \in R^{m \times (n-x)}$$

$$S_{a\mu}^{\leftarrow} = S_{a\mu}^{\rightarrow} \frac{\sum_{i} A_{ia}^{new} X_{i\mu} / (A^{new} S)_{i\mu}}{\sum_{k} A_{ka}^{new}} \quad (7)$$

$$S' = [s'_1, s'_2, \dots, s'_N], \quad s' \in R^x$$

표 1 잡음 데이터 경우의 분류성능 비교

Class	Method- I		Method- II	
	correct	incorrect	correct	incorrect
Speech (Male)	30	0	30	0
Speech (Female)	13	17	25	5
Music	10	0	9	1
Total	53	17	64	6

이 과정은 잡음이나 다른 신호들에 대한 또 다른 기저를 허용하고, 특징 S' 은 그런 기저와는 관계가 없는 것이므로, 특징 S' 을 잡음에 강인하게 만든다. 이는 귀속의 hair-cell을 n 개라고 가정하고 학습을 모델을 학습시키고, 그 중에서 x 개의 기저는 분류에 유용한 hair-cell이라고 생각할 수 있고, $n-x$ 개의 기저는 공통된 잡음에 대한 hair-cell이라 생각할 수 있을 것이다. 방법 I과 같이 그저 선택된 기저만을 사용한다면 공통된 잡음 및 소리에 대한 부호화가 NMF의 비 음수화 제약으로 인해, 선택된 $n-x$ 개의 기저에 가중되어 전체적인 특징을 흐리게 될 것이다. 따라서 식 (6)과 (7)에서는 질의 신호에 대해 공통된 잡음 및 신호에 대한 기저역시 갱신식으로 표현하여 이를 보완한다. 표 1은 3-부류에 대한 분류 실험으로 모든 실험 데이터에 잡음을 추가하여, 이 특징의 잡음 강인성에 대한 실험을 보여준다. 이 표에서 분류 성능은 각각 비교되고, 제안한 방법 II가 더욱 잡음에 강인함을 볼 수 있다.

4. HMM 분류기

은닉 마르코프 모델은 초기 상태 분포 π_i 상태 변화 행렬 T_{ij} 관찰 밀집 함수 $b_j(o)$ 의 세 가지로 구성된다. 연속 HMM은 $b_j(o)$ 를 각 상태가 $B_j = \{\mu_j, K_j\}$ 로 주어지는 평균이 μ_j 이고 공분산이 K_j 인 다변수 가우시안 분포로 설정한다. 따라서 은닉 마르코프 모델 $\lambda = \{T_j, B_j, \pi_j\}$ 로 표기한다[14].

분류를 위해서 관찰된 데이터가 주어졌을 때 각 모델의 확률을 재는 유사도와 가장 비슷한 상태의 시퀀스 $I = \{i_1, i_2, \dots, i_T\}$ 가 관찰된 데이터 O 와 모델 파라미터 λ 가 주어졌을 때 측정된다. HMM 분류기는 가장 큰 유사도를 나타내는 모델을 N 개의 경쟁 모델 중에서 골라낸다.

$$N^* = \arg \{ \max_{1 \leq j \leq N} P(O, I | \lambda_j) \} \quad (8)$$

5. 실험 및 결과

실험 데이터는 음성은 TIMIT 데이터베이스, 음악을 위해서는 상업적 CD들, 그리고 악기와 외부 소리들에 대해서는 웹에서 다운로드한 샘플을 사용한다. 사용된 소리들의 길이는 5에서 15초이다. 데이터 집합은 40%의 학습 데이터와 60%의 테스트 데이터로 나뉜다.

실험에서는 모든 소리 샘플들은 8KHz로 재표본 되고, 25ms 길이의 해밍(Hamming) 윈도우와 15ms 길이의 중첩을 갖는 STFT(Short Time Fourier Transform)를 이용하여 계산된다. 계산된 스펙트로그램은 100ms의 길이로 50ms씩 이동하면서 시간 축에 따라 나뉘어져서 데이터 행렬로 구성된다. NMF 갱신 법칙인 수식 (3)과 (4)에 의해서 구성된 데이터는 150개의 기저 벡터를 계산해낸다. 찾아진 150개의 기저는 분별력 측정 방법 (5)에 의해서 정렬된다. 정렬된 기저 중 90%가 유지되도록 경계값을 설정해서 113개의 정렬된 기저 벡터를 선택한다. 선택된 113개의 기저 벡터는 방법 II를 사용하여 부호화 변수 (특징)를 추정하기 위해 사용된다.

시스템의 테스트를 위해서는, 10가지 부류의 오디오 데이터와 5개의 은닉 상태를 갖는 HMM 분류기를 이용하여 보편적인 최대 유사도 측정법을 이용하여 10개의 은닉 마르코프 모델을 학습시킨다. 추출된 특징은 각 HMM의 입력으로 들어가고, 가장 큰 유사도를 갖는 모델 (8)을 이용하여 선택된다(그림 4).

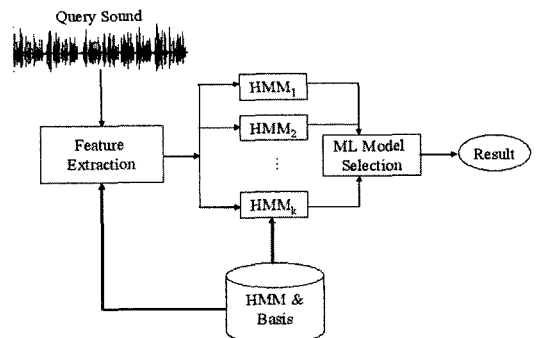


그림 4 소리 분류 시스템

우리는 잡음 있는 환경에서의 음성/음악 구분 실험과, 일반적인 소리 구분의 두 가지 실험을 하였다. 잡음 환경에서의 음성/음악 구분 실험에서는 HMM 분류기는 깨끗한 신호를 사용하여 학습되고, 성능은 5dB의 백색 잡음이 들어가 있는 테스트 데이터를 이용하여 측정된다(표 1의 방법 II 참조).

일반 소리 분류를 위해서는 잡음은 고려하지 않았으며, 분류 성능은 같은 학습 및 테스트 데이터를 사용하여 ICA 기반 방법과 비교하였다. 올바른 분류는 Hit로 계수되며, 틀린 분류는 Missed로 계수된다. 전체적인 성능은 총 126개의 테스트 데이터에서 올바른 분류의 비율로 표시하였다. 결과적으로 두 가지 방법이 모두 좋은 성능을 보이지만, NMF 기반 방법이 ICA 기반 방법보다 더욱 좋은 성능을 보임을 확인할 수 있다. 이 결과는 비음수 제약이 오디오 신호의 정확한 특징을 추출하는데 유용하다는 것을 보여준다. 특히 Violin과 Dog-Bark 소리에서 가장 큰 성능의 차이를 보이는 것은 제안한 방법을 이용하여 추출해낸 특징이 ICA에 기반한 방법보다 더욱 시간 축 및 주파수 축으로 국지화된 특징을 보인다는 것을 증명한다 하겠다. 다시 말해서 제안한 방법이 ICA에 기반한 방법으로 확인할 수 없는 특징을 더욱 국지화된 표현 방법을 적용함으로써 추출해 낼 수 있다는 것을 보여준다.

표 2 제안한 방법과 M.Casey의 ICA방법과의 분류 성능 비교

Class	Proposed		M.Casey's	
	Hit	Missed	Hit	Missed
Speech (Male)	30	0	30	0
Speech (Female)	30	0	28	2
Music	9	1	9	1
DogBark	9	0	2	7
Cello	10	0	9	1
Flute	9	1	9	1
Violin	7	0	2	5
Footsteps	9	0	8	1
Applause	3	2	2	3
Trumpet	4	2	5	1
Totals	120	6	104	22
Performance	95.24%		82.54%	

6. 결론

우리는 여기서 NMF에 기반을 둔, 스펙트럴 소리에서 특징을 추출하는 방법에 대해 서술했다. ICA에 기반을 둔 방법과 비교했을 때, NMF에 기반을 둔 방법은 부분 기반 표현법과 가까운 보다 국지화된 특징을 보여준다. 또한 기저와 관찰된 데이터가 주어졌을 때, 부호화

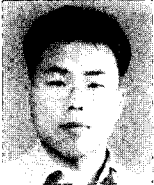
행렬을 추정하는 방법들에 대해 서술했다. NMF 기저 벡터의 국지적인 특징과 더불어 제안한 기저 선택 방법은 분류 성능을 향상시켰다. 보편적인 HMM 분류기를 이용하여 제안한 방법은 ICA에 기반을 둔 방법보다 훨씬 뛰어난 성능을 보였다.

참고 문헌

- [1] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, Vol. 40, pp. 788-791, Oct. 1999.
- [2] M. Casey, "Sound classification and similarity tools," in *Introduction to MPEG-7: Multimedia Content Description Language*, B. S. Manjunath, P. Salembier, and T. Sikora, Eds. John Wiley & Sons, Inc., 2001.
- [3] M. Casey, "Reduced-rank spectra and minimum-entropy priors as consistent and reliable cues for generalized sound recognition," in *Proc. Workshop on Consistent and Reliable Acoustic Cues for Sound Analysis*, Eurospeech, Aalborg, Denmark, 2001.
- [4] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, Inc., 2001.
- [5] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, John Wiley & Sons, Inc., 2002.
- [6] A. Bell and T. Sejnowski, "Learning the higher-order structure of a natural sound," *Network: Computation in Neural Systems*, Vol. 7, pp. 261-266, 1996.
- [7] M. S. Lewicki, "Efficient coding of natural sounds," *Nature Neuroscience*, Vol. 5, No. 4, pp. 356-363, 2002.
- [8] K. P. Kording, P. Konig, and D. J. Klein, "Learning of sparse auditory receptive fields," in *Proc. IJCNN*, Honolulu, Hawaii, 2002.
- [9] B. A. Olshausen and D. J. Field, "Sparse coding with an overcomplete basis set: A strategy employed by V1," *Vision Research*, Vol. 37, pp. 3311-3325, 1997.
- [10] D. A. Depireux, J. Z. Simon, D. J. Klein, and S. A. Shamma, "Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex," *J. Neuro-physiology*, Vol. 85, pp. 1220-1234, 2001.
- [11] S. Shamma, "On the role of space and time in auditory processing," *TRENDS in Cognitive Science*, Vol. 5, No. 8, pp. 340-348, 2001.
- [12] M. S. Gazzaniga, R. B. Ivry, and G. R. Mangun, *Cognitive Neuroscience: The Biology of the Mind*, W. W. Norton & Company, New York, 2001.
- [13] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Advances in*

Neural Information Processing Systems, Vol. 13, 2001.

[14] L. R. Rabiner and B. H. Juang, "An introduction to hidden Markov models," IEEE trans, Acoustics, Speech and Signal Processing Magazine, Vol 3, pp. 4-16, 1986.



조 용 춘

1994년~2002년 전남대학교 컴퓨터공학과 학사. 2002년~2004년 포항공과대학교 컴퓨터공학과 석사. 2004년~현재 삼성전자 DM총괄 DM연구소 Audio Lab. 연구원. 관심분야는 Audio Signal Processing, Neural Network



최 승 진

1987년 서울대학교 전기공학과(공학사)
1989년 서울대학교 전기공학과(공학석사)
1996년 University of Notre Dame(공학박사). 1997년 일본 RIKEN Frontier 연구원. 1998년~2003년 일본 RIKEN Brain Science Institute, Invited Senior Research Fellow. 1997년~2001년 충북대학교 전기전자공학부 조교수. 2001년~현재 포항공과대학교 컴퓨터공학과 조교수. 관심분야는 Statistical machine learning, Graphical models, Bayesian statistics, Information theory, Neural networks

방 승 양

정보과학회논문지 : 소프트웨어 및 응용
제 31 권 제 1 호 참조