

청각 기억 게임을 위한 리듬 유사도 측정 기술

Measurement of Rhythmic Similarity for Auditory Memory Game

김 주 완*, 이 세 원*, 박 호 종*
(Juwan Kim*, Sewon Lee*, Hochong Park*)

*광운대학교 전자공학과
(접수일자: 2011년 2월 15일; 채택일자: 2011년 3월 29일)

본 논문에서는 청각 기억 게임을 위하여 두 개의 소리 파형을 비교하여 파형의 리듬 유사도를 정량적으로 측정하는 기술을 제안한다. 제안한 방법은 두 입력 파형에 대하여 에너지 변화, 에너지 피크의 지속 시간, 음색 등을 분석하여 각 파형에 포함된 비트 위치를 검출하고, 두 파형의 템포 차이와 비트 수의 차이를 보상하는 과정을 통하여 두 파형의 리듬 벡터를 각각 정의한다. 다음, 두 리듬 벡터 사이의 차이와 비트 수의 차이를 적용하여 두 입력 파형의 리듬 유사도를 정량적으로 표현하는 식을 정의한다. 제안한 방법으로 측정한 리듬 유사도와 주관적 청취 평가로 측정한 리듬 유사도를 비교하였으며, 두 방법에 의한 리듬 유사도가 상관도 0.86을 가지는 것을 확인하였다.

핵심용어: 리듬 유사도, 비트 검출, 리듬 패턴

투고분야: 음악음향 및 음향심리 분야 (8.6)

In this paper, a method for measuring rhythmic similarity between two sound signals for auditory memory game is proposed. The proposed method analyzes energy fluctuation, the temporal duration of energy peak, the timbre of two signals, and detects beat positions for each signal. Then, it determines the rhythm vector after compensating a difference in tempo and the number of beats between two signals. Finally, a method for rhythmic similarity measurement is defined as a function of the dissimilarity between two rhythm vectors and a difference in the number of beats. The rhythmic similarity measured by the proposed method and that by the subjective listening test are compared, and the correlation of 0.86 between two results is achieved.

Keywords: Rhythmic similarity, Beat detection, Rhythm pattern

ASK subject classification: Musical Acoustics and Psychoacoustics (8.6)

I. 서론

사회가 고령화됨에 따라 고령자들에게 나타나는 인지 능력, 기억능력, 학습능력, 운동능력 등의 광범위한 기능의 저하를 보상하고 해당 능력을 향상시키기 위한 다양한 헬스 케어 산업이 발전하고 있다 [1-2]. 특히, 오락성을 추가하여 지루하지 않고 흥미를 가지면서 기능 저하를 방지하고 해당 능력을 향상시키려는 기능성 게임이 개발되고 있다 [3]. 기능성 게임은 게임 진행을 통하여 다양한 기능을 측정하고, 능력 향상을 위하여 특별히 설계된 동작을 반복 수행하여 해당 기능을 집중적으로 자극하고 반복적으로 활용하게 함으로써 해당 능력을 향상시키도록 한다.

본 논문에서는 고령자들의 청각적 인지, 기억 및 재현 능력을 측정하고 해당 기능을 향상시키는 청각 기억 게임을 위한 핵심 기술을 개발한다. 게임 사용자는 스네어 (snare) 드럼과 베이스 (bass) 드럼 소리로 이루어진 입력 소리를 청취하고, 소리가 표현하는 리듬 패턴을 인지하고 기억한다. 다음, 사용자는 기억한 리듬 패턴을 게임기의 입력 장치 (예로, 드럼 패드)를 통해 재현하여 게임기로 입력시킨다. 게임기는 원 소리 파형과 사용자가 재현한 소리 파형을 비교 분석하고, 두 개의 파형이 나타내는 리듬 패턴 사이의 유사도를 수치화하여 게임 사용자의 청각 인지, 기억 및 재현 능력을 종합적으로 측정한다. 또한, 게임기는 다양한 종류와 난이도의 리듬 패턴을 사용한 반복 수행을 통하여 청각 인지 기능과 기억 기능을 자극하여 청각 인지 및 기억 능력을 향상시키고, 기억한 리듬 패턴을 정확하게 재현하는 능력을 향상시켜 원하는

행동을 정확하게 수행하는 운동 능력을 향상시킨다. 이와 같은 청각 기억 게임을 구현하기 위하여 두 개의 파형에 포함된 리듬 패턴을 비교하여 리듬 유사도를 정량적으로 측정하는 것이 핵심이며, 본 논문에서는 리듬 패턴의 특성을 결정하는 파라미터를 정의하고 이를 비교하여 두 리듬의 유사도를 정량화하는 방법을 개발한다.

리듬 패턴을 비교하는 기술은 이전부터 많은 연구자에 의하여 연구되어져 왔다. 음악의 피치 정보와 리듬 패턴 정보를 추출하여 멀티미디어 신호를 검색하고 분류하는 기술이 개발되었고 [4], 음악 이론을 바탕으로 리듬 형식의 분석을 위하여 리듬 패턴들을 비교하고 멜로디의 알고리즘 및 기하학적 리듬 패턴을 분석하는 기술이 개발되었다 [5]. 또한 리듬 패턴의 차이를 해밍 거리 측정법, 기하학적 벡터 거리 측정법, 간격 비율 거리 측정법, 교환 거리 측정법 등으로 측정하여 각각의 성능을 비교하였다 [6]. 그러나 이와 같은 기존의 리듬 패턴 비교 기술들을 본 논문에서 개발하는 청각 기억 게임에 적용하기에는 문제점들이 있다. 즉, 기존의 기술들은 여러 개의 리듬 패턴들 사이의 유사한 정도를 판단하여 “상대적” 순위를 측정하며, 이는 데이터베이스에서 원하는 정보를 검색하는 응용분야에는 적합하지만, 청각 기억 게임과 같이 두 리듬 사이의 “절대적” 유사도를 정량화 하는 응용에는 적합하지 않다. 따라서 두 개의 파형이 나타내는 리듬 사이의 유사도를 정량적으로 측정하는 새로운 기술이 필요하다.

본 논문에서는 악보의 음표 정보가 아니라 실제 드럼으로 연주된 오디오 신호 파형을 이용하여 리듬 유사도를 측정하는 기술을 제안한다. 제안한 기술은 파형으로부터 리듬 비교를 위한 특징 파라미터를 추출하고, 추출된 파라미터를 처리하여 벡터화 하고, 리듬 벡터를 비교하여 유사도를 측정하는 기술로 구성된다. 또한, 드럼 사운드의 음색을 구별하기 위한 추가 동작을 포함하여 드럼 사운드의 리듬 추출 성능을 향상시켰다 [7-8]. 제안한 방법으로 측정된 리듬 유사도와 주관적 청취 평가로 측정된 리듬 유사도를 비교하였으며, 두 방법에 의한 리듬 유사도가 상관도 0.86을 가지는 것을 확인하였다.

II. 제안한 리듬 유사도 측정 기술

2.1. 개요

본 논문에서 제안하는 리듬 유사도 측정 기술의 전체 구조는 그림 1과 같다. 두 개의 파형이 입력되면 각 파

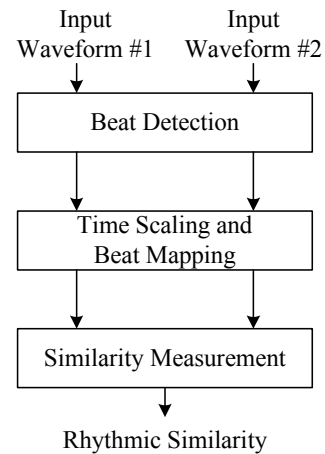


그림 1. 제안하는 리듬 유사도 측정 기술의 구조
Fig. 1. Block diagram of the proposed method.

표 1. 리듬 유사도의 정량적 측정 기준
Table 1. Reference scale for rhythmic similarity.

Comparison Scale	
100	Equal
80	Very similar
60	Similar
40	Slightly similar
20	Different

형의 특성 분석을 통하여 리듬 패턴을 구성하는 비트 (beat)의 시간적 위치를 검출한다. 다음, 두 리듬의 템포 (tempo) 차이를 보상하고 비트 사이의 상대적 간격만으로 표현되는 리듬 패턴을 검출하기 위하여 시간 스케일링을 실시하고, 그 결과를 리듬 벡터로 정의한다. 만일 두 파형의 비트 수가 일치하지 않을 경우에는 비트 수를 일치시키는 비트 매핑 (beat mapping) 과정을 추가로 수행하여 리듬 벡터를 정의한다. 마지막으로 두 리듬 벡터를 비교하여 두 파형의 리듬 유사도를 구한다. 유사도를 정량적으로 측정하기 위하여 유사도에 대한 기준이 필요하며, 본 논문에서는 표 1의 기준에 따라 유사도를 정량적으로 측정하고, 성능 평가도 표 1의 기준에 따른 주관적 평가와 비교하여 실시한다.

2.2. 비트 검출

드럼으로 연주된 리듬 패턴은 드럼 비트의 시간적 위치에 의하여 결정되므로, 드럼 비트 위치를 검출하는 과정이 먼저 수행된다. 드럼 비트는 시간 축에서 에너지의 급격한 변화로 쉽게 추출되므로 파형에서 에너지 피크 발생 시점 검출을 통하여 해당 비트를 검출한다. 입력 파형을 10 ms 단위의 프레임으로 구분하고, 실험에 의해 결정된

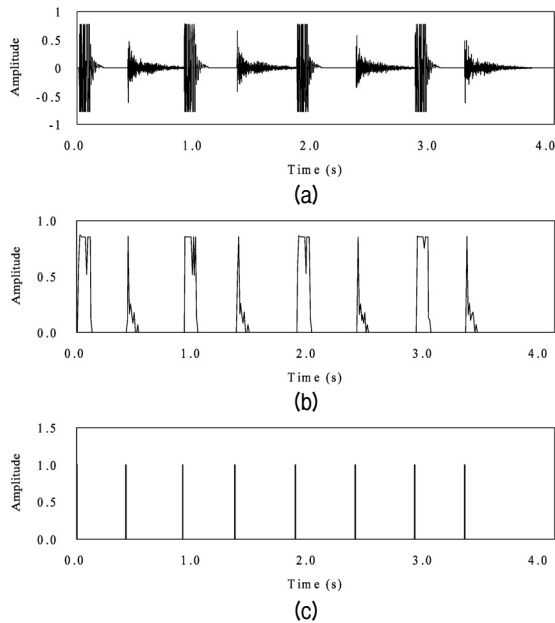


그림 2. 입력 파형에서 비트를 검출하는 예. (a) 입력 파형 (b) 프레임별 에너지 (c) 검출된 비트 위치
 Fig. 2. Example of beat detection from input waveform. (a) Input waveform (b) Frame energy (c) Detected beat positions.

임계값을 이용하여 해당 프레임의 에너지가 임계값보다 크면 비트가 있는 것으로 판정한다. 일반적으로 하나의 드럼 비트는 연속된 다수의 프레임에 걸쳐 큰 에너지를 가지고 그에 따라 연속 프레임에서 비트로 인식되므로, 임계값을 넘는 최초 프레임의 시간 정보만을 검출하여 비트 위치로 정의한다. 그림 2가 입력 파형에서 비트를 검출하는 예를 보여준다. (a)는 입력 파형이고, (b)는 프레임별 에너지이고, (c)는 최종적으로 검출된 비트 위치이다.

비트 검출 과정에는 miss detection 현상 (실제 비트가 있지만 비트로 검출되지 않은 경우)과 false alarm 현상 (비트가 없지만 비트로 검출되는 경우) 이 자주 발생한다. 이러한 현상들은 울림통 안에서 공명 현상이 발생하는 스네어와 베이스 드럼의 긴 여운 특성 때문에 발생하며, 단순히 에너지의 크기 분석만으로는 해결이 불가능하므로 에너지 피크의 지속 시간과 ZCR (zero crossing rate)을 통한 음색 구분을 사용하여 해결하도록 한다. 베이스 드럼의 경우 그림 3에 표시된 영역과 같이 한 번의 비트에 해당하는 구간 중간에 에너지가 급격히 작아지는 구간이 존재하며, 이에 따라 여러 번의 비트로 인식하는 false alarm 문제가 발생한다. 이 구간은 에너지 피크 발생 후 0.1초 내에서만 발생되며 약기의 특성상 0.1초 사이에 같은 음색의 연속적인 비트가 발생되지 않는다. 따라

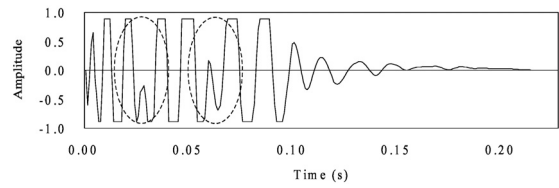


그림 3. 베이스 드럼의 파형
 Fig. 3. Waveform of base drum.

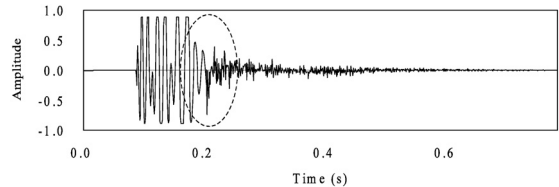


그림 4. 연속된 베이스와 스네어 드럼의 파형
 Fig. 4. Waveform of consecutive base drum and snare drum.

서 0.1초의 구간 내에서는 추가적인 비트를 검출하지 않는 방법으로 false alarm 문제를 해결할 수 있다.

드럼 연주 시 여러 종류의 드럼을 동시에 사용하면 짧은 시간 안에 서로 다른 음색의 비트 발생이 가능하다. 예로, 스네어와 베이스 드럼이 짧은 시간 내에 연속적으로 타격이 이루어지는 경우가 있는데, 이러한 연주 특성이 miss detection 문제를 발생시키는 원인이 된다. 그림 4는 처음 베이스 드럼의 비트가 발생하고 약 0.1 초 후에 스네어 드럼의 비트가 발생했을 때의 파형이다. 표시된 영역과 같이 짧은 간격 사이에 발생된 두 개의 서로 다른 비트는 에너지 피크와 지속 시간만으로는 정확한 검출이 어려우므로 음색 (timbre)에 대한 구별이 반드시 필요하다. 본 논문에서는 베이스와 스네어 드럼의 음색 구분을 위하여 식 (1)으로 정의되는 ZCR을 사용하고, $Sign(n)$ 은 n 번째 샘플의 부호 (± 1)를 나타내고 N 은 프레임의 샘플 수이다.

$$ZCR = \frac{1}{2N} \sum_{n=0}^{N-1} |Sign(n) - Sign(n-1)| \quad (1)$$

ZCR은 대략적인 주파수 측정을 통하여 음색 차이를 보여주고, 스네어 드럼과 베이스 드럼의 ZCR는 큰 차이를 가진다. 따라서 프레임 진행에서 ZCR 변화가 실험을 통하여 결정된 임계값 이상을 가지면 드럼 종류가 변경된 것으로 판단하여 짧은 시간 이내에서도 새로운 비트로 검출한다.

비트 검출 후에는 각 비트 위치 정보를 리듬 벡터로 표현한다. 입력 파형에서 K 개의 비트를 검출하고, 각

비트 위치가 초 단위로 b_k , $k=0, \dots, K-1$ 라 할 때, 리듬 벡터는 $B=[b_0, b_1, \dots, b_{K-1}]$ 로 표현된다.

2.3. 스케일링

전문적으로 훈련된 연주가가 아니면 리듬 패턴을 정확히 인지하더라도 리듬을 재현할 때 템포 (tempo) 오류가 일어나게 된다. 따라서 청취한 리듬과 기억 후 재현한 리듬 사이의 절대적 템포를 정확히 일치시키는 것은 매우 어렵다. 그러나 리듬 패턴은 각 비트 사이의 상대적 위치로부터 결정되므로, 리듬 유사도 측정에서 두 리듬 사이의 절대 템포의 차이를 배제하기 위하여 시간 축 스케일링을 적용하여 동일한 템포의 리듬 벡터로 변환한다. 이 때, 모든 벡터의 첫 성분은 0.0, 마지막 성분은 1.0 이 되도록 정규화 한다.

리듬 벡터 $B=[b_0, b_1, \dots, b_{K-1}]$ 의 스케일링은 식 (2)에 따라 b_k 를 새로운 비트 위치 b'_k 로 변환하는 과정으로 이루어진다.

$$b'_k = (b_k - b_0) \frac{1}{b_{K-1} - b_0}, \quad k=0, \dots, K-1 \quad (2)$$

이렇게 하면 새로운 비트 사이의 간격은 $(b'_{k+1} - b'_k) = (b_{k+1} - b_k) \frac{1}{b_{K-1} - b_0}$ 이 되어 모든 비트 간격이 일정 비율로 변하여 리듬 패턴의 변화는 없게 된다. 또한, 항상 $b'_0 = 0.0$, $b'_{K-1} = 1.0$ 을 만족하고, 스케일링 후의 새로운 리듬 벡터를 $B' = [b'_0, b'_1, \dots, b'_{K-1}]$ 로 정의한다.

두 리듬 벡터의 비트 수 (즉, 벡터의 길이)가 동일하면 각 벡터를 스케일링 하고, 스케일링 된 두 벡터를 비교하여 리듬 유사도를 측정한다. 만일 두 벡터의 비트 수가 서로 다르다면 다음 절에서 설명하는 비트 매핑 (beat mapping) 과정과 스케일링 과정을 동시에 수행하여 여러 개의 벡터를 정의하고 최종 리듬 유사도를 측정한다.

2.4. 비트 매핑

청각 기억 게임에서 사용자가 청취한 리듬 파형과 재현한 리듬 파형의 비트 수가 일치하지 않는 경우가 자주 발생하며, 이 경우에 두 리듬 벡터를 그대로 스케일링 하여 비교하면 잘못된 리듬 유사도를 측정하게 된다. 예로, 게임 사용자가 게임에서 제시한 리듬 패턴을 정확히 기억하여 입력한 후 마지막에 실수로 잘못된 비트를 하나 더 입력한 경우, 리듬 유사도는 매우 높고 단지 마지막의 잘못된 비트로 인하여 유사도가 감소되는 결과가 나와야

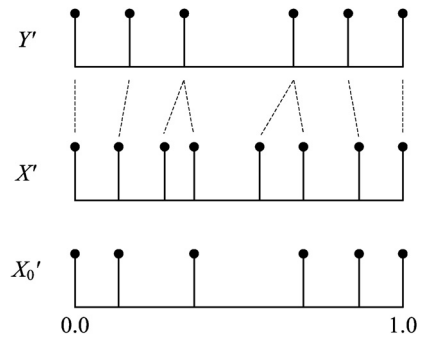


그림 5. 비트 매핑에서 비트 수를 일치시키는 과정의 예
Fig. 5. Example of beat removal process in beat mapping.

한다. 그러나 만일 비트 수 차이를 무시하고 그대로 스케일링 하여 두 리듬 벡터를 비교하면 마지막 오류 비트로 인하여 모든 비트의 위치가 잘못 되어 매우 낮은 유사도로 판정하는 오류를 범한다. 이와 같이 비트 수가 일치하지 않을 경우에 나타나는 문제점을 해결하기 위하여, 비트가 많은 벡터에서 비트를 선택적으로 제거하여 두 리듬 벡터의 비트 수를 일치시키는 비트 매핑 과정을 거친다. 물론, 비트 수 차이는 최종 리듬 유사도 점수에 반영된다.

비교할 두 리듬 벡터 중에서 비트 수가 많은 리듬 벡터를 X , 비트 수가 적은 리듬 벡터를 Y , 두 리듬 벡터의 비트 수를 각각 J 와 I , $J > I$ 라 한다. 이 때, 제시된 리듬 벡터와 재현된 리듬 벡터는 구분하지 않고, 단지 비트 수에 따라 X 와 Y 를 정의한다. 본 논문에서는 이해를 쉽게 하기 위하여 $J=8$, $I=6$ 으로 고정하여 X 와 Y 사이의 비트 매핑 과정을 설명하며, 이 과정은 매우 쉽게 임의의 J 와 I 로 일반화 될 수 있다. 먼저, 벡터 X 에서 비트를 한 개 또는 두 개를 제거하여 새로운 벡터들을 생성한다. 이 때, 임의의 위치에서 비트를 제거하면 리듬 패턴이 크게 변형되기 때문에 시작 부분과 끝 부분에서만 비트를 제거한다. 즉, X 에서 첫 비트를 제거하여 X_1 를 생성하고, 마지막 비트를 제거하여 X_2 를 생성하며, X_1 와 X_2 는 각각 7개의 비트를 가진다. 또한 X 에서 첫 두 비트를 제거하여 X_3 를 생성하고, 마지막 두 비트를 제거하여 X_4 를 생성하고, 마지막으로 X 에서 첫 비트와 마지막 비트를 제거하여 X_5 를 생성한다. X_3 , X_4 , X_5 는 각각 6개의 비트를 가진다.

X , X_1 , X_2 의 경우에는 여전히 Y 의 비트 개수와 차이를 가지므로 다음 과정을 통하여 Y 와 비트 수를 일치시킨다. X 를 스케일링 하여 임시 벡터 X' 를 구하고 이를 Y' (Y 에 스케일링을 적용한 결과)와 비교하여 각 비트에 대하여 X' 와 Y' 사이에서 시간 위치가 가장 어긋나는 비트 두 개를 X' 에서 제거하여 최종적으로 6개의 비트를

가지는 X_0' 을 정의한다. 그림 5가 이 과정의 예를 도식적으로 보여준다. 각 그림은 시간 축에서 비트의 위치를 나타내며, X' 와 Y' 의 각 비트의 위치를 비교하여 X' 와 Y' 에서 위치가 가장 근접한 비트끼리 비트 쌍을 구성하면, X' 의 3번째 비트와 5번째 비트가 Y' 비트와 가장 어긋나는 것을 알 수 있고, 그에 따라 3번째와 5번째 비트를 제거하여 6개의 비트를 가지는 X_0' 을 생성한다.

동일한 방법으로 X_1 과 X_2 를 스케일링하고 Y' 와 비교하고, 시간 위치가 가장 어긋나는 한 개의 비트를 X_1 , X_2 에서 각각 제거하여 최종적으로 6개의 비트를 가지는 X_1' 과 X_2' 를 각각 생성한다. 반면에 X_3 , X_4 , X_5 의 경우는 Y 와 비트 수가 일치하기 때문에 스케일링만 적용하여 X_3' , X_4' , X_5' 를 생성한다. 이상의 비트 매핑과 스케일링 과정을 통하여 최종 생성된 6개의 벡터, X_0' , X_1' , X_2' , X_3' , X_4' , X_5' 과 Y' 의 유사도를 각각 측정하고, 가장 높은 유사도를 X 와 Y 사이의 리듬 유사도로 최종 결정한다.

2.5. 리듬 유사도 측정

X 에 스케일링과 비트 매핑 과정을 적용하여 X_j' , $j=0,1,\dots$ 를 구한 후, X_j' 와 Y' 를 비교하여 가장 유사도가 높은 경우를 선택하여 최종 리듬 유사도로 출력한다. 만일 X 와 Y 의 비트 수가 동일하면 $j=0$ 으로 한정한다. 먼저, X_j' 와 Y' 의 각 i 번째 성분을 $(\cdot)_i$ 로 정의할 때, 식 (3)과 같이 각 벡터가 정의하는 비트 간격의 상대적 비율의 평균을 구하여 X_j' 와 Y' 사이의 리듬 일치 정도를 수치화 한다. 여기서 I 는 X_j' 와 Y' 의 비트 수이다. S 는 두 리듬의 일치 정도를 0과 1 사이의 값으로 표현하고, 만일 리듬 패턴이 정확히 일치하면 $S=1.0$ 이 된다.

$$S = \frac{\sum_{i=1}^{I-1} \min\{(X_j')_i - (X_j')_{i-1}, (Y')_i - (Y')_{i-1}\}}{\sum_{i=1}^{I-1} \max\{(X_j')_i - (X_j')_{i-1}, (Y')_i - (Y')_{i-1}\}} \quad (3)$$

마지막으로 X 와 Y 의 비트 수 차이를 적용하여 식 (4)와 같이 최종 리듬 유사도를 정의한다.

$$\text{리듬 유사도} = 100 \times S^{\frac{7}{2}} \times \left(\frac{I}{J}\right)^{\frac{9}{2}} \quad (4)$$

J 와 I 는 비트 매핑 과정 전 최초 X 와 Y 각각의 비트 수이며, 앞에서 정의한 것 같이 $J \geq I$ 이다. 청취 실험을 통하여, 동일 비트 수를 가지면서 두 리듬의 비트 위치에

차이를 가지는 경우보다 두 리듬의 비트 수 차이가 더 큰 오류로 인지되는 것을 확인하였으며, 그에 따라 상대적으로 더 민감한 비트 수의 오류에 더 큰 가중치를 부여하였고, 각 가중치 값은 실험을 통하여 정하였다.

III. 성능 평가

성능 평가는 제안한 기술로 측정된 리듬 유사도와 주관적 청취 평가로 측정된 리듬 유사도 사이의 상관관계 (correlation)를 측정하는 방법으로 진행하였다. 평가에 사용된 DB는 전자 키보드의 드럼 음원을 통하여 생성하였으며, 5-7초 길이를 가지는 총 5개의 기준 리듬 파형을 생성하였다. 다음, 각 기준 리듬 파형에 대하여 4가지의 변형된 리듬 파형을 생성하고, 이를 기준 리듬 파형과 비교하여 각각의 유사도를 측정하였다. 따라서 총 20 가지의 테스트 조합이 구성되며, 제안한 방법으로 각 조합에 대한 리듬 유사도를 측정하였다.

주관적 청취 평가에는 20~30대 남성과 여성을 포함하여 총 12명이 참여하였으며, 총 20 테스트 조합에 대하여 각각 기준 리듬 파형과 변형된 리듬 파형을 연속적으로 청취하고 두 리듬 파형 사이의 주관적 유사도를 표 1의 기준에 따라 평가하였다. 음원 청취는 잡음 영향이 없는 환경에서 헤드폰을 사용하여 진행하였다.

제안한 방법에 의한 유사도 측정 결과와 주관적 청취 평가 결과 사이의 상관관계를 식 (5)를 통하여 측정하였으며, $r=0.86$ 을 얻었다.

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} \quad (5)$$

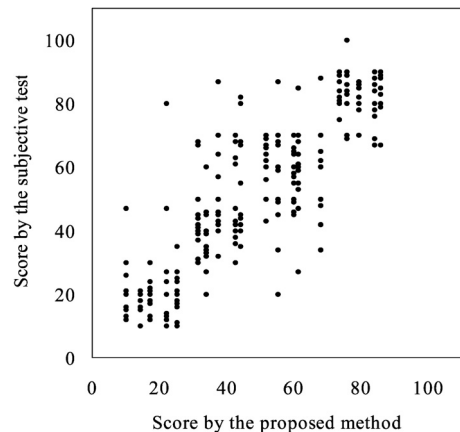


그림 6. 제안된 방법과 청취 평가로 측정된 리듬 유사도 분포
Fig. 6. Distribution of rhythmic similarity measured by the proposed method and subjective listening test.

여기서 x_i, y_i 는 각 테스트 조합에 대하여 제안한 방법으로 측정된 점수와 청취 평가로 측정된 점수를 각각 나타내고, 각각의 평균이 \bar{x} 와 \bar{y} 이다.

그림 6은 제안한 방법으로 측정된 유사도와 주관적 청취 평가로 측정된 유사도의 상호 분포를 보여준다. 두 개의 서로 다른 방법으로 측정된 유사도 점수가 대각선 영역에 집중되어 높은 상관관계를 가지는 것을 알 수 있다. 물론, 평가자 개인마다 유사도에 대한 주관적 판단 기준이 다르기 때문에 제안한 방법과 완전 일치하는 결과를 얻는 것은 불가능하다. 이상과 같이, 상관도 $r=0.86$ 과 그림 6의 분포를 통하여 제안한 방법으로 측정된 리듬 유사도가 실제 사람이 듣고 느끼는 리듬의 유사 정도와 비교적 일치하는 것을 확인하였다.

IV. 결론

본 논문에서는 청각 기억 게임을 위한 핵심 기술인 두 소리 파형 사이의 리듬 유사도 측정 기술을 제안하였다. 드럼으로 연주된 두 파형의 시간적 에너지 변화, 에너지 피크의 지속 시간, 음색 등을 분석하여 파형에서의 비트 위치를 각각 검출하고, 두 리듬 사이의 템포 차이를 보상하여 상대적 리듬 정보만을 추출하기 위하여 시간 스케일링을 실시하였다. 다음, 두 파형에서의 비트 수의 차이를 처리하는 과정을 거쳐 최종 리듬 유사도를 측정하는 방법을 개발하였다. 제안한 리듬 유사도 측정 방법의 성능을 주관적 청취 평가 결과와의 비교를 통하여 검증하였으며, 두 방법에 의한 리듬 유사도가 0.86의 상관도를 가지는 것을 확인하였다. 제안한 기술은 청각 인지, 기억 및 재현 능력을 향상시키는 게임의 핵심 기술로 사용될 수 있고, 향후 다양한 타악기의 리듬 패턴에 대한 유사도 측정 기술로 발전시켜 음악 신호처리 분야에서 복잡한 리듬 패턴 인식에 적용하여 리듬 훈련을 위한 도구로도 활용될 수 있다.

감사의 글

본 연구는 2010년 광운대학교 교내 연구비 지원으로 진행되었습니다.

참고 문헌

1. 김경, 권대규, "감각 및 운동기능 증진을 위한 헬스케어 로봇 기술동향", *대한전기학회 전기의 세계*, 제28권, 제7호, 20-28쪽, 2009.
2. 김형지, 유미, 진혜련, 유재성, 권대규, "복합감각 기능증진 개념의 헬스케어 목욕시스템 개발", *한국정밀공학회 2010년도 춘계학술대회 논문집 (하)*, 1395-1396쪽, 2010.
3. 김강수, 한용희, 정원범, 이영호, 강정훈, 최흥호, 문치웅, "다채널 근전도 신호를 이용한 체감형 게임 인터페이스 개발", *한국게임학회 논문지*, 제10권, 제5호, 65-73쪽, 2010.
4. Edward J. Coyle and Ilya Shmulevich, "A system for Machine Recognition of Music Patterns", *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1998.
5. Godfried T. Toussaint, "Algorithmic, Geometric, and Combinatorial Problems in Computational Music Theory", *X Encuentros de Geometria Computacional*, pp. 101-107, 2003.
6. Godfried T. Toussaint, "A Comparison of Rhythmic Similarity Measures", *5th International Conference on Music Information Retrieval*, 2004.
7. Adam Tindale, Ajay Kapur, and Ichiro Fujinaga, "Towards Timbre Recognition of Percussive sounds," *International Computer Music Conference (ICMC)*, 2004.
8. Adan Tindale, Ajay Kapur, George Tzanetakis and Ichiro Fujinaga, "Retrieval of Percussion Gestures Using Timbre Classification Techniques," *International Conference on Music Information Retrieval*, 2004.

저자 약력

●김 주 완 (Juwon Kim)

2009년 2월: 대구대학교 전자시스템공학과 (공학사)
 2011년 2월: 광운대학교 전자공학과 (공학석사)
 2011년 1월 ~ 현재: 코리아일레콤 연구원
 ※ 주관심 분야: 오디오 신호처리, 통신 신호처리

●이 세 원 (Sewon Lee)

2000년 2월: 광운대학교 전자공학과 (공학사)
 2002년 2월: 광운대학교 전자공학과 (공학석사)
 2004년 3월 ~ 2007년 8월: 다산네트웍스 책임연구원
 2007년 9월 ~ 현재: 광운대학교 전자공학과 박사과정
 ※ 주관심 분야: 음성/오디오 신호처리, 오디오 정보 추출

●박 호 중 (Hochong Park)

1986년 2월: 서울대학교 전자공학과 (공학사)
 1987년 12월: Univ. of Wisconsin-Madison, 전기공학과 (M.S.)
 1993년 5월: Univ. of Wisconsin-Madison, 전기공학과 (Ph.D.)
 1993년 9월 ~ 1997년 8월: 삼성전자 선임연구원
 1997년 9월 ~ 현재: 광운대학교 전자공학과 교수
 ※ 주관심 분야: 음성/오디오 신호처리, 통신 신호처리, 영상 신호처리