

스펙트럼 변조를 이용한 청각정보의 촉감재현 가능성 연구

Feasibility Study on Audio-Tactile Display via Spectral Modulation

곽현구¹, 김희국^{1,✉}, 정주노¹, 강대임², 박연규², 구민모³
Hyun-Koo Kwak¹, Whee Kuk Kim^{1,✉}, Juno Chung¹, Dae Im Kang², Yon Kyu Park² and Min-Mo Koo³

¹ 고려대학교 제어계측공학과 (Department of Control and Instrumentation Engineering, Korea Univ.)

² 한국표준과학연구원 (Division of Physical Metrology, Korea Research Institute of Standards and Science)

³ 고려대학교 지혜과학연구소 (Wisdom Science Center, Korea Univ.)

✉ Corresponding author: wheekuk@korea.ac.kr, Tel: 041-860-1443

Manuscript received: 2010.3.23 / Revised: 2010.11.1 / Accepted: 2011.1.25

Various approaches directly using vibrations of speakers have been suggested to effectively display the aural information such as the music to the hearing-impaired or the deaf. However, in these approaches, the human can't sense the frequency information over the maximum perceivable vibro-tactile frequency (around 1kHz). Therefore, in this study, an approach via spectral modulation of compressing the high frequency audio information into perceivable vibro-tactile frequency domain and outputting the modulated signals through the designated speakers is proposed. Then it is shown, through simulations of using Short-Time Fourier Transform (STFT) with Hanning windows and through preliminary experiments of using the vibro-tactile display testbed which is built and interfaced with a notebook PC, that the modulated signal of a natural sound composing sounds of a frog, a bird, and a water stream could produce the noise-free signal suitable enough for vibro-tactile speakers without causing significant interfering disturbances. Lastly, for three different combinations of information provided to the subject, that is, i) with only video image, ii) with video image along with the modulated vibro-tactile stimuli as proposed in this study to the forearm of the subject, and iii) with video image along with full audio information, the effects to the human sense of reality and his emotion to given audio-video clips including various sounds and images are investigated and compared. It is shown from results of those experiments that the proposed method of providing modulated vibro-tactile stimuli along with the video images to the human has very high feasibility to transmit pseudo-aural sense to the human.

Key Words: Audio-tactile Display (청각-진동 촉감 재현), Spectral Modulation (스펙트럼 변조), Audio Frequency (가청 주파수), Perceivable Vibrotactile Frequency (가촉 주파수)

1. 서론

사람의 지각기능이 외부로부터의 모든 입력정보를 효과적으로 처리하지 못하는 경우 사람의 가용한 다른 지각기능을 활용하는(sensory substitution) 또는 교차감각을 이용한 정보전달(cross-modal

display) 방안들에 관한 많은 연구가 진행되고 있다.¹⁻⁴ 인공보철의 삽입이 어려워 청각을 통한 정보 전달이 불가능한 청각 장애인에게도 주변상황이나 음악소리 등과 관련된 청각정보를 제공하기 위한 노력으로서 시각, 촉각 등 청각이외의 다양한 형태의 감각기능을 통하여 청각 정보를 전달하고자

다양한 방안들이 시도되고 있다.^{3,5,6}

사람의 촉각은 신체기관의 모든 피부를 통하여 다양한 형태의 감각적인 촉각정보를 얻을 수 있을 뿐 만 아니라 반응속도의 측면에서(0.2 초내외) 비교할 때 청각이나 시각과 크게 다르지 않은 정보 처리 능력을 가지고 있어 교차감각의 전달매체로서 큰 잠재성 및 가능성이 있다 하겠다.

그러나, 사람의 청각과 촉각에서 감지할 수 있는 가청 주파수 영역과 가촉 주파수(perceivable vibro-tactile frequency) 영역은 각각 20Hz~20kHz 그리고 10Hz~1kHz로서 약 20 배 정도의 차이가 존재한다. 또한, 임의의 주파수에서 주파수의 변화를 감지할 수 있는 가를 나타내는 주파수 민감도(Weber's Fraction= $\Delta f/f$)는 청각의 경우는 0.3%로 매우 민감하지만 촉각의 경우 전박부의 경우는 10%에서 40%, 손가락의 경우는 20%에서 60%까지로 매우 둔감한 것으로 보고되고 있다.⁷ 따라서, European Telecommunications Standards Institute 는 진동촉감을 "tactile icon (tacton)"과 같은 추상적인 신호패턴으로 사용하는 데 있어 구분 가능한 주파수 단계는 7 단계까지 그리고 촉각에 가해지는 변위에 대해 명확히 구분 가능한 단계는 15 단계까지로 제한할 것을 권장하고 있다.

또한, 촉각의 경우 동일한 촉각 자극에 대해 빨리 둔감해지는 신호 적응도(adaption rate) 특성, 신체 피부의 위치에 따라 다른 그러나 손가락과 같은 민감한 부분을 제외한 신체 대부분의 경우 매우 낮은 분해능 특성을 나타내는 자극 위치 민감도(spatial acuity), 그리고 주어진 자극의 동시처리 기능의 한계 등이 존재한다. 따라서, 이러한 한계를 극복하고 촉각을 활용하여 박자(rhythm), 가락(melody), 음색(timber), 그리고 넓은 음역(register)을 통해 느껴지는 청감을 효과적으로 재현하기 위해서는 많은 진전이 필요하다 하겠다.

소리-진동 촉감 재현(audio-tactile display)에 관련된 연구로서 주변의 청각 정보를 신체의 여러 적절한 위치에 부착된 다수의 진동모터의 진동을 통하여 제공하거나, 추상적 개념이나 또는 주변상황에 관한 주의 상기의 목적으로 촉감 아이콘(tactons) (신호형태, 주파수, 리듬, 그리고 소리세기 등에 근거한 다수의 진동모터 어레이에 설정된 진동신호 패턴)을 활용하는 방법이 제시되어 왔다.^{9,11} 그러나, 촉감 아이콘을 활용하는 방법은 주위환경 변화에서 감지되는 청각 정보를 설정된 추상적 개념의 아이콘을 통하여 전달하기 때문에 정보소통

에 있어서는 어느 정도의 효과가 인정되고 있으나 청감을 효과적으로 전달하기에는 한계가 있다.

본 연구자들이 파악하기로는 현재까지 청각 장애인에게 청감을 제시하고 자 하는 연구는 거의 없었으나 최근 들어 음악에서 연상될 수 있는 청감을 촉감을 통하여 사람에게 재현시키고자 하는 노력의 일환으로 여러 개의 스피커를 활용한 방법이 Karam¹¹ 외와 Nanayakkara¹² 외 등에 의해 시도되고 있다. 그러나, 제시된 방법들의 경우 현실적으로 사람의 가촉 주파수 영역이 대략 10Hz에서 1kHz의 범위로 제한되기 때문에 스피커를 통하여 전달된 1kHz 이상의 높은 주파수 대역에 속하는 음악정보는 거의 인식되지 않게 되어 많은 음악정보의 손실이 수반되는 단점이 존재한다.

따라서, 본 연구에서는 넓은 주파수 대역의 청각 정보를 가촉 주파수 영역내로 변조한 후 변조된 음악정보를 스피커를 통하여 출력하는 방안을 제시하고 자연음을 대상으로 한 모의실험을 통하여 원 자연음의 고주파수 정보가 가촉 주파수영역내로 효과적으로 변조될 수 있도록 하였다. 그리고 변조된 촉각 진동자극이 청각 장애인에게 적용되는 경우 어느 정도의 유사한 청감을 전달할 수 있는 지에 대한 가능성을 확인하기 위한 기초 연구로서 정상인을 대상으로 하되 i) 본 연구에서 제시된 방법에 의해 변조된 청각정보에 해당하는 진동촉각 정보와 청각 정보와 연관된 영상정보가 제공되는 경우, ii) 영상정보만이 제공되는 경우, 그리고 iii) 청각 정보 및 영상정보가 모두 제공되는 경우인 세 가지 다른 실험조건(이들은 각각 청각 장애인에게 변조된 진동 촉감 정보를 제공하는 경우, 시각 정보 만에 의존하는 청각 장애인의 경우, 그리고 시각 및 청각 정보를 접하는 정상인의 경우와 유사한 조건임)에 대하여 14 가지의 다른 유형의 소리를 내포하는 환경에 대해 어느 정도의 실제감과 정서감을 느낄 수 있는지에 관한 비교 분석을 수행하였다.

2. 청각신호의 진동촉감 재현 시스템

Fig. 1은 청각 신호를 입력 받아 스펙트럼 변조를 수행한 후 변조된 신호를 진동 촉감 재현기기를 통해 사람의 피부에 제공하는 청각신호 진동 촉감 재현 시스템의 흐름도이다. 그림에서와 같이 진동촉감재현 시스템은 크게 두 가지 단계(스펙트럼 변조 단계와 진동 촉감 재현 단계)로 나누어

질 수 있으며 입력된 청각신호가 진동촉감을 통하여 사람이 실제로 느끼는 청감에 유사한 느낌을 가질 수 있도록 하기 위해서 효과적으로 개발 및 조합되어야 한다.

스펙트럼 변조 단계: 입력된 청각정보의 손실을 최소화 하면서 진동 촉감 재현 기기에 적합한 주파수 영역의 신호로 변조한다. 특히, 입력된 청각신호를 압축 변조하되 변조과정에서 발생할 수 있는 불연속성으로 인한 잡음이 최소화된 신호를 도출한다.

진동 촉감 재현 단계: 청각 장애인이 스펙트럼 변조 단계를 통해 변조된 신호를 효과적으로 감지할 수 있도록 변조된 신호에 대하여 주파수에 따른 세기를 변조하고 주파수 영역을 분할하여 해당 진동 스피커로 출력한다. 이 단계에서는 효과적인 진동 촉감 재현 기기 구현 및 제시 방법 연구 그리고 청각 장애인의 감성 및 인지 효능 평가를 통한 검증 작업이 요구된다.

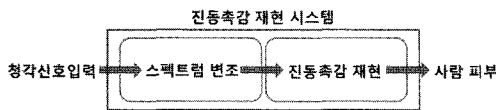


Fig. 1 Information flow diagram of vibro-tactile display system

따라서, 본 연구에서는 주파수 변조 단계에 관련된 내용으로서 입력된 청각 신호의 스펙트럼 변조와 주파수 분할방식을 제시한 후 그 적용 가능성을 조사하기 위하여 다수의 스피커를 조합한 진동 촉감 시험대를 제작하여 제시된 스펙트럼 변조 및 분할 방식(spectral contraction and division method:SCDM)에 대한 기초 실험을 수행하였다. 그리고 진동 촉감 재현 단계에 속하는 내용으로서는 제시된 SCDM 방식에 의한 신호가 다수의 스피커 진동을 통하여 청각 장애인에게 전달되었을 때 감지 효능 평가에 관한 한 기초 연구로서 정상인을 대상으로 하여 제시된 스펙트럼 압축 변조 및 분할 방식의 가능성을 확인하였다.

3. 스펙트럼 변조 방안

우리가 흔히 듣는 타 악기(percussion)와 금관 악기(brass)들에 대한 바탕 음의 기본 주파수(fundamental frequency)의 범위는 대부분 1kHz 영역

내에 존재하지만 목관악기(woodwinds)나 현악기(strings)에 대한 기본 주파수의 범위는 3kHz 로부터 5kHz 까지 걸쳐 있다. 파이프 오르간(pipe organ)의 경우는 바탕 음의 기본 주파수가 15kHz 까지 달한다. 한편, 악기의 음색을 나타내는 배음(harmonics)의 대부분은 1kHz 에서 최대 17kHz 까지의 영역에 걸쳐 있다. 이러한 일반적인 음악 정보가 스피커를 통하여 그대로 전달되는 경우 가측 주파수보다 큰 1kHz 이상의 청각 정보(거의 모든 악기의 음색 정보)는 사람이 인식하지 못하므로 그대로 버려지게 됨을 알 수 있다.

이러한 단점을 극복하기 위해 본 연구에서는 모든 주파수영역의 청각정보를 가측 주파수 영역으로 압축 변조한후 변조된 주파수 영역을 분할하여 여러 개의 스피커를 통하여 진동촉감을 재현하는 방안을 제시한다. 이 방식에서는 청각신호의 스펙트럼에서 전체 또는 원하는 주파수 이하의 정보를 가측 주파수 영역안(10Hz~1kHz)으로 다음과 같이 압축 변조한다.

$$f_{tactile} = c_1 f_{aural} \tag{1}$$

여기서 $f_{tactile}$, f_{aural} 그리고 c_1 은 각각 진동 촉감 주파수, 입력 청각신호 주파수 그리고 압축비를 나타낸다. 이러한 압축 변조 방식은 상대적으로 낮은 청각-진동 촉감 변환 분해능을 가지게 되어 제한적인 청감의 전달 효능을 가질 것으로 보이지만 넓은 주파수 영역대의 청각 정보를 전달할 수 있는 이점이 있다.

4. STFT 를 이용한 스펙트럼 변조

Fig. 2 는 스펙트럼의 압축변조 모의실험에서 사용된 STFT 알고리즘¹³의 요약도이다. 이 알고리즘은 Phase Vocoder(PV)와 유사한 기법으로서 스펙트럼의 압축변조과정에서 발생하는 잡음을 최소화하기 위하여 Hanning window 가 추가된 방법이다. 적용된 스펙트럼 변조기법을 요약하면 다음과 같다:

단계 1) 입력된 청각 신호를 설정된 크기의 이층의 Hanning window 를 통하여 각각 통과시킨 후 통과된 각 window 내의 데이터들에 대해 Fast Fourier Transform(FFT) 변환을 통하여 스펙트럼을 구한다.

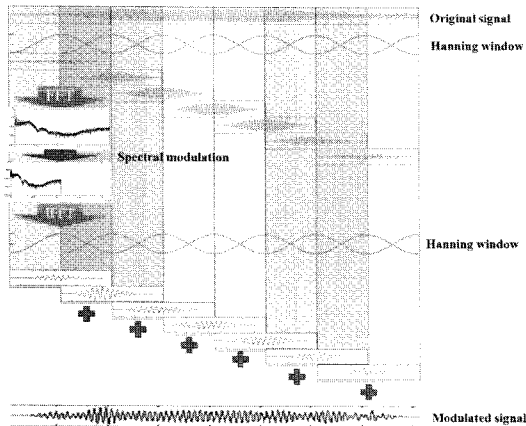


Fig. 2 STFT algorithm employing Hanning windows

단계 2) 각 스펙트럼을 원하는 형태로 변조한다. Fig. 3 은 압축 변조비가 1/2 인 경우에 변조된 스펙트럼과 위상각을 산출하는 과정을 나타낸다. 여러가지 방법이 있을 수 있으나 편의상 그림에서와 같이 본 연구에서는 원하는 샘플 속도의 1/2 (변조에서 요구되는 스펙트럼 압축비에 해당)에 해당하는 샘플 수로서 스펙트럼을 구한 후 이들을 주파수 상에서 압축하는 방법을 활용하였으며 이러한 과정은 물리적으로는 음조를 낮추는(pitch shifting down) 효과를 나타낸다.

단계 3) 압축 변조된 각 스펙트럼을 Inverse FFT(IFFT) 변환시킨 후 또 다시 Hanning window 를 통과시킴으로서 위상각의 변화로 인한 창과 창사이의 불연속점들을 최소화 한 후 모든 신호들을 수합하여 스피커를 통하여 출력한다.

단계 1 에서 입력신호에 사각형 모양의 창(window) 대신에 Hanning window 를 사용한 것은 단계 2 에서와 같이 임의의 스펙트럼 변조가 있는 경우 단계 3 에서 IFFT 를 통해 변환된 시간영역의 신호에서 창과 창사이의 경계부근에서 불연속점(glitches)이 발생되며 이로 인하여 스피커에서 상당한 크기의 잡음(clicks)을 발생되게 되는데 이러한 잡음을 최소화 하는 효과를 얻기 위함이다.

그리고 단계 3 에서 사용된 Hanning window 의 경우는 임의의 스펙트럼 변조에서 발생하는 주파수 및 위상각의 변화로 인해 상당한 크기의 잡음이 각 창의 경계점에서 나타나는 데 이러한 잡음은 진동 촉감 재현에 상당한 영향을 끼치므로 이를 최소화 하기 위한 목적으로 윈 청각정보로부터의 약간의 손실을 감수하며 사용되었다. 실제로, STFT

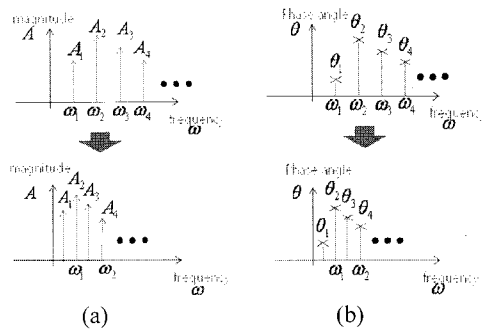


Fig. 3 (a) Magnitude modulation with compression ratio of 1/2, (b) Phase angle modulation with 1/2 compression ratio

과정내에서 IFFT 된 신호들에 대해 Hanning window 가 사용되었을 때 시뮬레이션 및 스피커를 통한 소리 출력을 통하여 창의 경계점 부근에서 발생하는 잡음이 거의 모두 제거되는 것을 확인하였다.

한편, 이상의 STFT 알고리즘을 적용하는 데 있어 입력신호의 샘플속도는 44.1kHz 로 설정하였으며 각 창은 4096 샘플 데이터로 구성하였다. 그리고 본 알고리즘과 관련된 제반 신호처리 과정은 PC 상의 Matlab 를 이용하였다.

5. 자연소리에 대한 모의실험

사람이 진동으로 감지할 수 있는 가측 주파수 영역은 1kHz 이하로 이 보다 높은 주파수의 소리를 포함하는 경우 많은 청각 정보가 진동촉각을

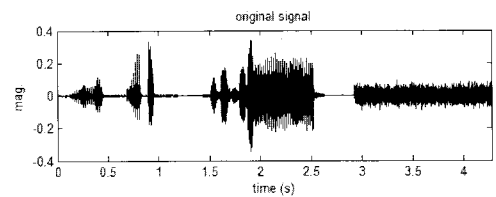


Fig. 4(a) Signal of natural sound formed with sounds of a frog, a bird and a water stream

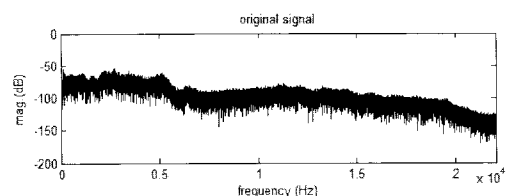


Fig. 4(b) Its magnitude spectrum

통하여 전달되지 못하게 된다. 이러한 청각 정보의 손실을 확인하기 위하여 본 모의 실험에서 사용된 청각 입력 정보는 자연음으로써 낮은 주파수 영역에 속하는 개구리 소리, 높은 주파수 영역에 속하는 새소리, 그리고 넓은 대역폭을 가지는 물소리로 구성하였다. Fig. 4(a)와 Fig. 4(b)는 각각 시간영역에서의 자연음과 전체 시간 데이터에 대한 스펙트럼을 나타낸다.

5.1 진동-촉감 재현 시험대

Fig. 5 는 진동 스피커를 이용한 진동-촉감 재현 시험 모듈이다. 이러한 진동-촉감 재현 시험 모듈은 Haptic chair 와 같은 의자형 진동 촉감 시스템에 설치되어 사람의 팔이나 다리 또는 등과 같은 다양한 위치에 부착될 수 있는 모듈의 형태로 설계되었으며 Fig. 5 에 보인 시험 모듈은 총 8 개의 진동 스피커로 구성되며 Velcro 에 고정되어 사람의 전박부(forearm)에 자연스럽게 접촉이 잘 될 수 있는 형태로 제작되었다. 이들 8 개의 진동 스피커는 각각 Fig. 6 에 보인 바와 같이 노트북 PC 의 상용 sound card 와 연결된 스피커 앰프를 통하여 제어된다. 그림에서와 같이 각 스피커 앰프는 스테레오 출력이 가능하여 두 개의 진동 스피커를 제어할 수 있다.

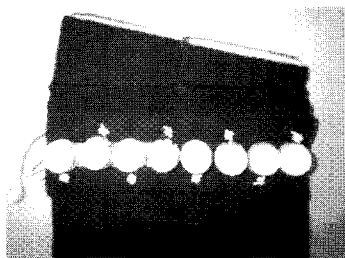


Fig. 5 Vibro-tactile display module composed of eight speakers attached on the Velcro

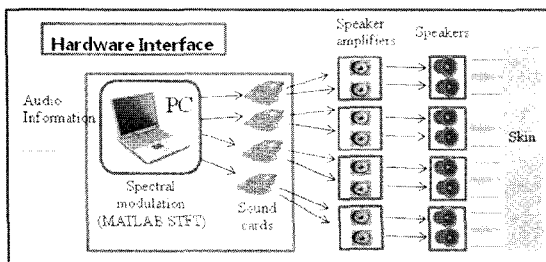


Fig. 6 Computer-speaker output interface

5.2 압축 변조에 관한 시뮬레이션

Fig. 7(a)와 Fig. 7(b) 그리고 Fig. 8(a)와 Fig. 8(b) 는 각각 개구리, 새 그리고 물소리로 구성된 자연음에서 1kHz 이하의 주파수 성분만을 취한 신호와 1kHz 이상의 모든 주파수 성분만을 취한 신호와

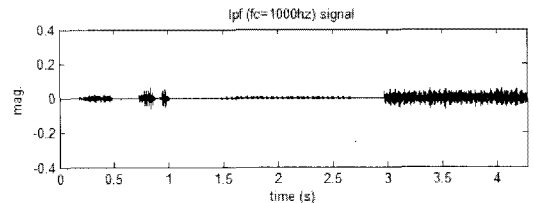


Fig. 7(a) Signal of the natural sound only with perceivable vibro-tactile frequency components (less than 1kHz)

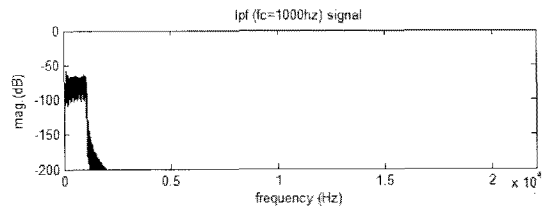


Fig. 7(b) Magnitude spectrum of signal of the natural sound only with perceivable vibro-tactile frequency components (less than 1kHz)

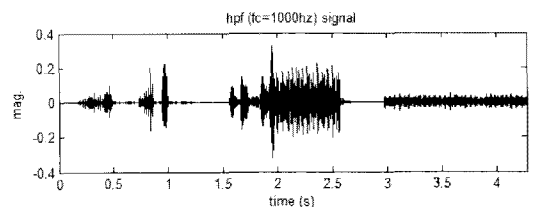


Fig. 8(a) Signal of the natural sound only with unperceivable vibro-tactile frequency components (over 1kHz)

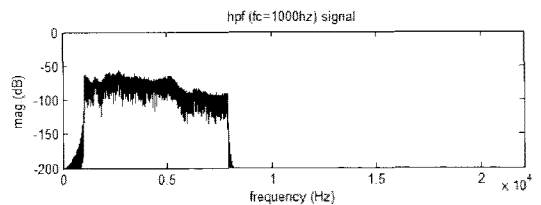


Fig. 8(b) Magnitude spectrum of the natural sound only with unperceivable vibro-tactile frequency components (over 1kHz)

스펙트럼이다. Fig. 8은 Fig. 4의 원 자연음 신호로부터 Fig. 7의 가속주파수 영역내의 주파수 정보를 제외한 모든 1kHz 이상의 손실된 청각 정보를 나타낸다. 실제로 Fig. 7(a)의 신호를 스피커를 통하여 듣는 경우 예상한 바와 같이 주로 1kHz 이하의 주파수에 걸쳐있는 개구리 소리는 어느 정도 원음에 가까운 소리가 유지되었지만 높은 주파수 영역에 해당하는 새소리는 거의 알아듣기가 어려울 정도로 큰 정보의 손실이 있음을 확인할 수 있었다.

Fig. 9(a)와 Fig. 9(b)는 원음을 1/4로 압축 변조한 후 1kHz 이상의 모든 주파수 성분을 제거한 신호 및 스펙트럼이다.

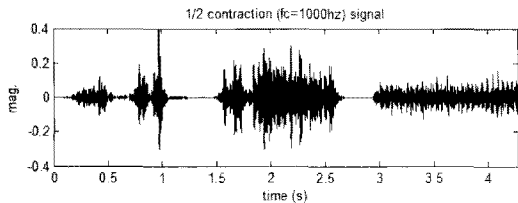


Fig. 9(a) Signal of the natural sound compressed with compression ratio of 1/4

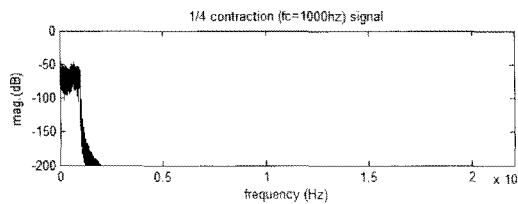


Fig. 9(b) Its magnitude spectrum

5.3 압축 변조된 신호의 세기 변조 및 주파수 영역분할

보다 효율적인 청각 재현 결과를 얻기 위해서는 진동 주파수에 따른 스펙트럼의 세기 변조와 주파수 영역의 분할에 대하여 보다 심층적인 연구가 필요하지만, 본 연구에서는 편의상, 압축 변조된 신호는 주어진 진동신호의 주파수에 따른 사람의 동일감(equal sensation) 특성⁸을 고려하여 얻어진 다음과 같은 패핑 함수에 따라 스펙트럼의 세기를 변조하였다.

$$f(x) = 1000 \times (0.0367x^6 - 0.4713x^5 + 2.4661x^4 - 6.6795x^3 + 9.8104x^2 - 7.3863x^1 + 2.2539) \quad (2)$$

여기서 x 는 주파수를 나타낸다.

그리고 진동 주파수에 따른 신호의 세기가 변조된 신호의 주파수 영역은 가장 민감한 진동 주파수 인지 영역으로 알려진 150Hz~250Hz 부근에서는 상대적으로 좁게, 그리고 나머지 영역은 넓게 조정하여 총 8개의 영역으로 분할하였으며 대략 2.5cm 정도의 간격으로 일렬로 놓인 진동 스피커를 통하여 신호가 출력되도록 하였다.

5.4 변조된 진동 촉감을 통한 유사 청각 재현 성능 평가 시스템

Fig. 10은 진동 촉감을 통한 유사 청각 재현 방식의 성능을 평가하기 위한 testbed 시스템으로써 8개의 진동 스피커가 일렬로 장착된 Velcro 촉각 재현 모듈, 청각 정보와 관련된 video clip을 제공하는 컴퓨터와 모니터, 그리고 의자와 필요시 제공되는 방음용 헤드폰으로 구성된다.



Fig. 10 Implemented vibro-tactile display test mockup

6. 스펙트럼 압축변조 방식의 인지 및 감성 실험

6.1 실험 조건 및 방법

본 실험에서는 앞에서 제시된 SCDM 방식을 활용하여 청각 장애인에게 어느 정도의 유사 청각을 제공할 수 있는가에 관한 기초 연구로서 일반 정상인을 대상으로 아래와 같은 세 가지의 실험 조건에 대해 실제감(생생감)과 정서감에 대한 반응을 각각 비교 분석하였다:

- i) 영상 정보(image)만을 제공,
 - ii) 영상정보와 SCD 진동촉감(vibro-tactile) 제공 (실험 대상자의 전박부),
 - iii) 영상 정보와 청각 정보 제공.
- 실제로 청각 장애인들의 경우 시각 정보는 정

상인들과 동일하게 입력되므로 영상 정보만이 입력되는 조건 i)는 청각 장애인의 조건과 유사하다고 볼 수 있다. 이와 같은 맥락에서 조건 ii)의 경우도 청각 장애인이 진동 촉감 재현 모듈을 장착한 경우로 볼 수 있다. 이와 같이 이들 두 실험 조건 i)과 ii)는 각각 청각 장애인 그리고 진동 촉감 재현 모듈을 장착한 청각 장애인과 유사하다고 볼 수 있다. 한편, iii)의 경우는 정상인의 상태를 나타낸다. 따라서, 이러한 세 가지의 실험 조건을 비교 분석함으로써 본 실험에서 목표로 하고 있는 청각 장애인의 유사 청각 재현 정도의 평가에 관한 유의한 결과를 확인하고자 하였다.

한편, 본 논문에서 제시된 SCDM 방식의 실시간 처리는 현재의 시스템으로는 너무 큰 전산하중이 요구되므로, 스펙트럼 변조 및 분할 데이터 처리는 실험에 앞서 수행되어 저장되었으며 이러한 저장된 데이터가 해당 실험 수행시 진동 촉감 재현 시스템에 입력되는 청각신호로서 제공되었다.

6.1.1 실험 대상자 선정

실험자 대상자로서는 대학교에 다니고 있는 학생들을 중심으로 구성하였다. 그러나 모든 실험 대상자는 세 가지의 실험 조건 중에서 한 가지 실험 조건에만 참여하도록 제한하였으며 또한, 실제감과 정서감 실험에서도 각각 한 가지 실험에만 참여하도록 제한하였다.

6.1.2 청각 정보의 선정

실험에 사용될 소리의 유형을 결정하기 위해 대전 청각 장애인 협회 회원 19 명을 대상으로 가장 듣고 싶어하는 소리 10 개를 선택하도록 하는 설문을 수행하였으며 얻어진 결과는 동물소리(14), 아기울음소리(13), 종소리(12), 경적소리(9), 계곡물소리(9), 사람 소리(8), 노래 소리(6), 초인종 소리(6), 기차 소리(5), 비소리(5), 웃음소리(5), 과일(물) 먹는 소리(4), 바람 소리(4), 악기소리(4), 비행기 소리(4) 등으로 요약된다. 여기서 괄호 안의 숫자는 빈도 수를 나타낸다. 이상과 같은 설문결과를 고려하여 모든 실험에서 공통적으로 제공되는 청각 정보로서는 아기 울음소리, 자동차 소리, 개구리 울음 소리, 팝콘소리, 종소리, 지하철 소리, 성인의 웃음소리, 폭포소리, 응원 합성, 고양이 울음소리, 개 울음소리, 사자 울음소리, 비바람 소리, 천둥소리인 총 14 가지의 소리를 선정한 후 각 유형의 소리와 관련된 시각 정보(video clip)를 수집하

여 활용하였다.

6.1.3 설문 유형

설문은 앞에서 선정한 14 가지의 다른 유형의 소리에 대한 질문으로써 주어진 상황에 관해 각각 얼마나 실제감 (또는 생생함) 또는 정서감을 느낄 수 있는가를 응답하도록 구성하였다. 예를 들면, 각 소리에 대한 실제감(또는 생생함)에 관한 설문은 “동영상속의 장면이 얼마나 생생하게 또는 사실적으로 느껴지는가?”로 모두 동일하지만 정서감에 관한 설문은 웃음소리에 관한 설문의 경우는 “웃고 있는 사람이 얼마나 즐거워 하는 것 같은가?”로 폭포소리에 관한 설문은 “폭포 소리가 얼마나 시끄럽게 느껴지는가?”와 같은 형태로 정서감이 내포되어 있도록 구성하였다.

6.1.4 실험 데이터 수집

실험조건 i) 즉, 영상 정보(image)만을 제공하는 실험조건에 대한 실제감과 정서감 인지 실험의 경우는 실험대상자들에게는 설문지가 먼저 배포되었으며 각 질문에 대해 5 초에 해당하는 영상 정보를 제공하고 난 후 응답하도록 하였다. 각 설문사이의 간격은 10~15 초 정도로 적절히 조정하였다.

영상정보(image)와 SCDM 진동 촉각 정보(vibrotactile)를 동시에 제공하는 실험조건 ii)에 대한 실제감 그리고 정서감 인지 실험의 경우는 실험대상자들에게는 5 초 동안 컴퓨터 모니터에 제시된 영상정보와 실험 대상자의 전박부에 제공되는 SCDM 진동촉각정보를 인지한 후 모니터에 제시된 질문에 따라 차례로 응답을 입력하도록 하였다.

그리고 영상정보와 청각정보를 동시에 제공하는 실험조건 iii)에 대한 실제감 그리고 정서감 인지 실험의 경우는 실험대상자는 5 초 동안 컴퓨터 모니터에 제시된 영상정보와 청각 정보를 제공한 후 모니터에 제시된 질문에 따라 차례로 응답을 입력하도록 하였다.

6.2 실험 결과 분석

6.2.1 소리 유형별 실험조건에 따른 실제감 인지에 대한 차이 검증

실험 대상자들을 14 종류의 소리와 관련하여 영상만 제공한 집단(이하 영상 집단, 20 명), 진동과 영상을 동시에 제공한 집단(이하 진동-영상 집단, 20 명), 소리와 영상을 동시에 제공한 집단(이하 소리-영상 집단, 20 명)으로 구분하여 각 소리를 실제

적인 것으로 인식하는 정도에 있어 집단 사이에 차이가 있는가를 비교하였다(Table 1). 분석 결과 영상 집단의 평균 실제감 점수는 51.20(±21.72), 진동-영상 집단은 65.95(±10.58), 소리-영상 집단은 67.30(±9.35)이었으며 이러한 집단들 사이의 평균 점수의 차이는 통계적으로 유의한 것으로 나타났다 (F=7.13, p=0.002).

Table 1 Test on differences among responses of reality perception and emotional perception under different experimental conditions

variable	image (a)	image+vibro-tactile (b)	sound+image (c)	significant test F(p)	post-hoc comparison (Scheffe)	
	mean ± SD	mean ± SD	mean ± SD		group comparison	p-value
reality perception ¹ (total score)	51.20 ±21.72	65.95 ±10.58	67.30 ±9.35	7.13** (0.002)	a*b	0.01**
					b*c	0.96
					a*c	0.005**
emotional perception ² (total score)	46.29 ±21.15	64.25 ±9.30	68.39 ±6.48	13.17** (<0.001)	a*b	0.001**
					b*c	0.64
					a*c	<0.001**

** : p ≤ 0.01

1:a (20 persons),b (20 persons),c (20 persons)

2:a (17 persons),b (20 persons),c (18 persons)

다음으로 세 집단 사이의 유의한 차이는 어떤 집단 사이의 차이로 인함인가를 분석하기 위해 2 집단씩 짝짓기를 하여 사후 검정을 시행한 결과 영상 집단과 진동-영상 집단 사이에는 실제감을 인식하는 정도에 있어 유의한 차이가 있는 것으로 나타난 반면(51.20±21.72 vs 65.95±10.58, p=0.01), 진동-영상 집단과 소리-영상 집단 사이의 차이는 유의하지 않은 것으로 나타났다(65.95±10.58 vs 67.30±9.35, p=0.96). 즉 영상 집단에 비해 진동-영상 집단이 소리 유형들에 대해 느끼는 실제감의 정도가 더 높은 것으로 나타난 반면, 진동-영상 집단과 소리-영상 집단 사이의 실제감의 인지 정도는 유의한 차이가 나지 않는 것으로 나타났다. 특정한 소리를 생생하게 느끼게 하는 면에서 볼 때 진동-영상은 영상만 제공한 경우에 비해 월등히 높은 효과를 보이는 한편, 소리-영상에 비해서 결

코 뒤떨어지지 않는 효과를 보이는 것으로 제시되었다.

한편, 각 소리별로 집단에 따른 실제감의 인지 정도를 살펴본 결과(Table 2), 진동-영상 집단의 평균 점수가 다른 집단들에 비해 가장 높은 것으로 나타난 소리의 유형은 아기 울음소리, 자동차 소리, 사자 울음소리, 총소리, 지하철 소리, 천둥 소리, 응원 소리 등이었다. 즉 본질적으로 소리에 강한 진동의 요소를 가지고 있는 유형의 경우는 영상만 제공하거나 소리-영상을 제공한 경우보다 진동-영상을 제공한 경우 더 실제감 있게 느끼는 것으로 나타났다. 반면 고양이 소리, 개구리 소리, 팝콘 소리 등과 같이 경쾌하거나 부드럽거나 소리의 크기가 작은 유형은 진동-영상 집단에 비해 소리-영상 집단의 평균 점수가 현저하게 높은 것으로 제시되었다.

Table 2 Technical statistics on reality perception for sound types under different experimental conditions

sound type	image	image+vibro-tactile	sound+image
	mean ± SD ¹	mean ± SD	mean ± SD
baby crying	1.00 ± 0.00	3.40 ± 1.67	3.00 ± 0.00
cat	4.45 ± 2.16	3.75 ± 1.68	5.70 ± 1.03
vehicle	3.95 ± 2.04	5.95 ± 1.19	4.45 ± 1.23
dog barking	3.70 ± 1.98	4.75 ± 1.55	5.40 ± 0.75
frog	3.75 ± 2.10	3.30 ± 2.00	4.70 ± 1.08
lion	3.95 ± 1.93	5.05 ± 1.40	5.00 ± 1.21
popping corn	3.75 ± 2.05	3.50 ± 1.73	4.80 ± 1.28
gun firing	4.40 ± 2.39	6.20 ± 1.06	4.95 ± 1.67
tempest	3.60 ± 2.14	5.05 ± 1.54	5.40 ± 1.23
subway train	4.40 ± 1.73	5.15 ± 1.35	4.55 ± 1.36
thunder	3.25 ± 1.80	6.05 ± 0.83	4.85 ± 1.04
laugh	3.45 ± 1.61	4.30 ± 1.84	5.10 ± 1.25
water fall	3.80 ± 1.85	4.80 ± 1.32	4.75 ± 1.57
rooting	3.75 ± 1.71	4.70 ± 1.56	4.65 ± 1.10

¹: standard deviation

6.2.2 소리 유형별 실험조건에 따른 정서감 인지에 대한 차이 검정

실제감 인지의 경우와 마찬가지로 실험 대상자들을 14 종류의 소리와 관련하여 영상만 제공한 집단(이하 영상 집단, 17 명), 진동과 영상을 동시

에 제공한 집단(이하 진동-영상 집단, 20 명), 소리와 영상을 동시에 제공한 집단(이하 소리-영상 집단, 18 명)으로 구분하여 각 소리와 연관되어 불러일으키는 정서감을 인지하는데 있어 집단 사이에 차이가 있는가를 비교하였다(Table 1). 분석 결과 영상 집단의 평균 정서감 인지 점수는 46.29(±21.15), 진동-영상 집단 64.25(±9.30), 소리-영상 집단 68.39(±6.48)이었으며 이러한 집단들 사이의 정서감 인지 점수의 차이는 통계적으로 유의한 것으로 나타났다(F=13.17, p<0.001).

다음으로 세 집단 사이의 유의한 차이는 어떤 집단 사이의 차이로 인한가를 분석하기 위해 2 집단씩 짝짓기를 하여 사후 검정을 시행한 결과 영상 집단과 진동-영상 집단 사이에는 정서감을 인지하는 정도에 있어 유의한 차이가 있는 것으로 나타난 반면(49.29±21.15 vs 64.25±9.30, p=0.001), 진동-영상 집단과 소리-영상 집단 사이의 차이는 유의하지 않은 것으로 나타났다(64.25±9.30 vs 68.39±6.48, p=0.64). 즉 영상 집단에 비해 진동-영상 집단이 소리 유형들에 대해 느끼는 정서감의 정도가 더 높은 것으로 나타난 반면, 진동-영상 집단과 소리-영상 집단 사이의 정서감의 인지 정도는 유의한 차이가 나지 않는 것으로 나타났다. 특정한 소리와 관련된 정서감을 불러일으키는 측면에서 볼 때 진동-영상은 영상만 제공한 경우에 비해 월등히 높은 효과를 보이는 한편, 소리-영상에 비해서 결코 뒤떨어지지 않는 효과를 보이는 것으로 제시되었다. 한편, 각 소리와 연관되어 느껴지는 정서감에 대한 집단별 평균 점수들을 살펴본 결과(Table 3), 진동-영상 집단의 평균 점수가 가장 높게 나타난 소리의 유형은 사자 울음소리, 총소리, 지하철 소리, 웃음 소리, 폭포 소리 등으로 실제감 인지와는 다소 다른 양상을 보였으나 진동-영상은 본질적으로 소리에 진동의 요소를 가지고 있는 소리 유형에 대해 그 소리와 연관된 정서를 불러일으키는데 더 효과적인 것으로 나타났다.

7. 논의

기존의 스피커를 이용한 청각정보의 진동촉각 재현 방식에서는 오직 가측 주파수 영역(1kHz)내에 있는 청각 정보만이 사람에게 전달되어 가측 주파수 영역밖에 속하는 고 주파수 영역의 청각 정보가 손실되는 단점이 존재한다. 따라서, 본 연

Table 3 Technical statistics on emotional perception for sound types under different experimental conditions

sound type	image	image+vibro-tactile	sound+image
	mean ± SD ¹	mean ± SD	mean ± SD
baby crying	1.00 ± 0.00	4.00 ± 1.56	6.17 ± 1.20
cat	4.06 ± 2.30	3.95 ± 1.32	5.22 ± 1.80
vehicle	3.94 ± 2.25	5.35 ± 1.23	5.39 ± 1.15
dog barking	3.94 ± 2.05	5.00 ± 1.30	6.22 ± 0.88
frog	3.53 ± 1.94	2.75 ± 1.62	3.17 ± 1.98
lion roaring	2.76 ± 1.60	5.20 ± 1.15	5.06 ± 0.80
popping corn	2.82 ± 1.81	2.75 ± 1.25	3.06 ± 1.47
gun firing	3.06 ± 1.75	4.70 ± 1.84	3.28 ± 1.74
tempest	4.00 ± 2.35	5.50 ± 1.19	6.28 ± 0.83
subway train	3.88 ± 1.87	4.85 ± 1.14	4.44 ± 1.42
thunder	3.12 ± 1.58	5.05 ± 1.47	5.44 ± 1.25
laugh	4.00 ± 2.03	5.95 ± 0.95	5.83 ± 0.86
water fall	2.71 ± 1.21	3.90 ± 0.85	2.94 ± 1.35
rooting	3.47 ± 1.74	5.30 ± 1.13	5.89 ± 1.08

¹: standard deviation

구에서는 보다 넓은 주파수 영역의 청각정보를 진동촉감을 통하여 효과적으로 전달할 수 있는 스펙트럼 압축변조 방식에 관하여 조사하였다. 구체적으로는 Matlab 환경에서 STFT 알고리즘을 활용한 스펙트럼 변조, 주파수에 따른 세기 변조 및 변조된 주파수 영역 분할, 노트북 PC 와 진동 스피커의 인터페이스, 총 8 개의 진동 스피커가 velcro 위에 일렬로 고정된 진동 촉각 재현 모듈 제작, 그리고 스펙트럼 변조된 소리에 대한 모의 실험을 수행하여 제시된 변조 방법을 통한 신호가 변조로 인한 잡음의 간섭 없이 안정적이며 효과적으로 넓은 주파수 영역의 청각정보를 가측 주파수 영역 안으로 압축 변조할 수 있음을 확인하였다.

그리고 제시된 스펙트럼 압축변조 및 분할 방식을 활용한 촉각 재현 방식이 어느 정도까지 청각 장애인에게 유사 청감을 재현할 수 있는가를 확인하기 위한 기초 연구로서 정상인을 대상으로 다양한 유형의 소리에 대한 환경 정보를 첫째, 영상 정보만을 제공하는 경우(청각 장애인의 환경), 둘째, 영상정보와 더불어 변조된 진동 촉감을 대상자의 전박부(forearm)에 제공하는 경우(제시된 방

안), 셋째, 영상정보와 청각 정보를 정상적으로 제공하는 경우(정상인의 환경)인 세 가지로 분류한 후 주어진 다양한 유형의 소리 환경에 대해 느끼는 실제감(생생함)과 정서감에 관한 반응을 비교 분석함으로써 두 번째 방식이 첫 번째 방식보다는 일반적인 정상인의 반응에 상당히 가깝다는 결과를 확인하였다.

본 논문에 이어 다음 단계의 연구로써는 i) 본 논문에서 제시된 청감-촉감 재현 방식 및 개발된 촉감 재현 모듈을 가지고 장애인을 대상으로 한 청감 재현의 효과를 검증하기 위한 실험 및 분석, ii) 보다 효과적인 변조 신호 처리 방식, 세기 변조 방식 그리고 주파수 분할 방식 등에 관한 연구, iii) 진동 촉감 뿐 만 아니라 열감과 같은 또 다른 형태의 촉감 자극을 추가하여 진동촉감과 함께 청감 재현의 효과를 향상시킬 수 있는 방안 등이 있으며 현재 이들에 관한 연구가 진행되고 있다.

후 기

이 논문은 2010년도 정부(교육과학기술부)의 재원으로 휴먼인지환경사업본부-신기술융합형 성장동력사업의 지원을 받아 수행된 연구임(No. 2010K001163).

참고문헌

- Sohn, M. S., Shin, H. S., Park, J. S. and Han, D. W., "Trends of Haptic Interface for Wearable Computers," Report of the Electronics and Telecommunication Research Institute of the Korea, Vol. 20, No. 5, pp. 149-155, 2005.
- Lindeman, R. W., Yanagida, Y., Noma, H. and Hosaka, K., "Wearable vibrotactile systems for virtual contact and information display," Virtual Reality, Vol. 9, No. 2, pp. 203-213, 2006.
- Hayward, V. and Maclean, K. E., "Do It Yourself Haptics : Part I," IEEE Robotics and Automation Magazine, Vol. 14 No. 4, pp. 88-104, 2007.
- Chouvardas, V. G., Miliou, A. N. and Hatalis, M. K., "Tactile displays : Overview and recent advances," Displays, Vol. 29, No. 3, pp. 185-194, 2008.
- Gunther, E., Davenport, G. and O'Modhrain, S., "Cutaneous Grooves: Composing for the Sense of Touch," Proceedings of Conf. on New Instruments for Musical Expression(NIME-02), pp.37-42, 2002.
- Park, Y.-K., Choi, I.-M., Kim, W.-K. and Kang, D.-I., "Eyesight and Hearing Assistant Technique using Multiple Sensory Display," J. of KSPE, Vol. 26, No. 11, pp. 7-11, 2009.
- Jones, L. A. and Sarter, N. B., "Tactile Displays : Guidance for Their Design and Application," Human Factors, Vol. 50, No. 1, pp. 90-111, 2008.
- Wada, C., Shoji, H. and Ifukube, T., "Development and evaluation of a tactile display for a tactile vocoder," Technology and Disability, Vol. 11. No. 3, pp. 151-159, 1999.
- Brown, L. M., Brewster, S. A. and Purchase, H. C., "A First Investigation into the Effectiveness of Tactons," Proceedings of the First Eurohaptics Conference and Symposium on Haptic Interfaces for virtual Environment and Teleoperator Systems, pp. 167-176, 2005.
- Ternes, T. and MacLean, K. E., "Designing Large Sets of Haptic Icons with Rhythm," Haptics: Perception, Devices and Scenarios, Vol. 5024, pp. 199-208, 2008.
- Karam, M., Russo, F., Branje, C., Price, E. and Fels, D. I., "Towards A Model Human Cochlea: Sensory substitution for crossmodal audio-tactile displays," Graphics Interface Conference, pp. 267-274, 2008.
- Nanayakkara, S., Taylor, E., Wyse, L. and Ong, S. H., "An Enhanced Musical Experience for the Deaf: Design and Evaluation of a Music Display and a Haptic Chair," Proceedings of the 27th International Conference on Human Factors in Computing Systems, pp. 337-346, 2009.
- Sethares, W. A., "Rhythm and Transforms," Springer, pp. 115-128, 2007.