

## TTS를 이용한 매장음원방송에서 고객의 인지도 향상을 위한 음향효과 연구

A Study on the Sound Effect for Improving Customer's Speech Recognition in  
the TTS-based Shop Music Broadcasting Service

강선미<sup>1)</sup> · 김현득<sup>2)</sup> · 장문수<sup>3)</sup>

Kang, Sunmee · Kim, Hyundeuc · Chang, Moonsoo

### ABSTRACT

This thesis describes the method for well voice announcement using the TTS(Text-To-Speech) technology in the shop music broadcasting service. Offering a high quality TTS sound service for each shop requires a great expense. According to a report on the architectural acoustics the room acoustic indexes such as reverberation time and early decay time are closely connected with a subjective awareness about acoustics. By using the result the customers will be able to recognize better the voice announcement by applying sound effect to speech files made by TTS. The result of an aural comprehension examination has shown better about almost all of the parameters by applying reverb effect to TTS sound.

**Keywords:** web music broadcasting, Cocktail Party Effect, Reverb Effect, Text-To- Speech

### 1. 서 론

인터넷의 빠른 발전으로 이를 이용한 많은 서비스들이 제공되고 있다. 그 중에 웹을 이용하여 매장에 방문한 고객들을 대상으로 음악 방송을 제공하는 서비스 시장도 빠르게 성장되고 있다[1]. 매장에 적절한 음악과 필요한 멘트를 사용함으로써 매장에 방문한 고객의 소비 촉진에 많은 영향을 미치게 된다. 멘트를 TTS(Text-To-Speech)로 제작하여 방송하는 시스템은 멘트 방송을 예약하거나 음악과 더불어 방송하는 등의 서비스가 가능하다. 따라서 사람이 직접 마이크로 멘트를 방송하는 것보다 다양한 서비스를 제공할 수 있다.

그러나 TTS 로 제작된 멘트는 TTS 의 성능에 따라 청취자에게 멘트의 내용이 정확하게 전달되지 않는 경우가 있다. 멘트의 음량을 높이는 것만으로는 청중의 만족도를 높일 수가 없다. 따라서 멘트에 적절한 효과를 주어 청취 만족도를 높여줌으로써 고객이 멘트에 좀 더 집중할 수 있는 방안을 제안하고자 한다.

### 2. TTS 를 이용한 멘트의 장단점

본 논문에 소개된 매장음원방송 서비스는 웹을 통하여 제공되고 있다. 각 매장의 분위기나 상황에 맞추어 실시간으로 멘트를 방송하려면 멘트를 TTS 를 이용하거나 매장에서 직접 마이크를 사용하는 방법이 가장 적합하다. TTS 를 이용하여 멘트를 작성할 경우 멘트 내용이 음원 파일로 작성되어 예약이 가능하다[2]. 즉, 서비스 시간에 직접 마이크로 녹음하지 않아도 되므로 방송 시간에 대한 제약을 받지 않는다. 그러나 녹음한 파일로 예약 기능을 제공하는 서비스를 구성할 경우 매장음원방송 서비스를 제공하는 서버에 파일을 업로드하는 불편함이 발생된다. 이러한 이유로 TTS 를 이용하여 멘트를 방송하는 시스템 구조가 웹에서 제공되는 매장음원방송 서비스에 적합하다.

1) 서경대학교 smkang@skuniv.ac.kr

2) 서경대학교 deempl@nate.com

3) 서경대학교 cosmos@skuniv.ac.kr 교신저자

(지원번호:)

접수일자: 2009 년 10 월 13 일

수정일자: 2009 년 10 월 30 일

게재결정: 2009 년 11 월 20 일

그러나 매장의 구성과 특성에 따라서 TTS 로 만들어진 멘트의 내용이 매장 내의 고객(청취자)에게 제대로 전달되지 못하는 경우도 생긴다. 사용자의 요구에 따라서 각 매장에 적합한 TTS 를 제공하는 것은 현실적으로 어려운 실정이다. 따라서 매장내의 고객들이 멘트 내용에 집중하도록 칵테일 파티 효과 등을 얻어낼 수 있도록 TTS 멘트에 음향 효과를 적용한다.

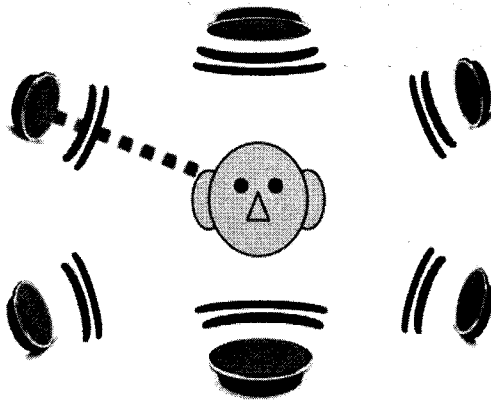


그림 1. 칵테일 파티 효과<sup>4)</sup>  
Figure 1. Cocktail Party Effect

건축음향에서 검증된 내용을 살펴보면 실내음향 지표는 청중의 위치에서의 잔향시간, 초기감쇠시간, 명료도, 요해도, 측면반사음 비율 양이상관도 등을 사용하여 공간의 목적에 적합하게 그 범위를 설정한다. 이러한 지표들은 청중의 잔향감, 명료도, 공간감 등 전체적인 음향의 만족도를 구성하는 주관적 인지도와 높은 관계가 있음이 알려져 있다[4]-[6]. 이는 청취자가 좋아하는 음향효과가 적용되면, 그 음향에 대해서 집중하여 듣게 되는 칵테일 파티 효과현상과 실내음향 지표와 연관되어 있음을 알 수 있다.

실내음향 지표의 내용들은 건축 실내 인테리어의 상태에 따른 변화도 적용되고 있으나, 본 논문에서는 방송음향 이외의 요소는 제외하고 실내음향 지표 중에서 잔향효과를 TTS 로 만들어진 멘트 음성파일에 적용하여 청중이 멘트의 인지도를 높이도록 제안한다.

**3. 잔향효과(Reverb Effect)**

청취자가 방송 시 멘트에 집중할 수 있게 하려면 청취자의 관심이 높은 멘트내용을 방송하거나 청취자가 좋아하는 효

4) 칵테일 파티 효과(Cocktail Party Effect) - 파티장, 나이트 클럽, 공사장 등 시끄러운 공간에서도 서로 대화가 가능한 이유는 자기에게 의미 있는 정보나 자신이 좋아하는 소리만을 선택적으로 받아들이는 선택적 지각(Selective perception)을 하기 때문이다. 이렇게 자신이 좋아하는 혹은 필요한 정보를 집중하여 듣는 현상을 칵테일 파티 효과라 한다[3].

과를 멘트에 적용하여야 한다. 후자의 경우를 살리기 위해 일반적으로 사람들이 선호하는 형태로 효과를 입혀 방송하는 방법을 선택한다.

일반적으로 사람들은 소리에 울림효과가 있는 것을 좋아한다[7]. 예를 들어 공항에서 멘트 방송을 할 때 약간의 울림 형태로 방송을 한다. 사람들은 자신의 이름을 부르지 않아도 그 방송에 집중하는 현상을 나타낸다. 그러나 이러한 울림효과가 너무 과하게 들어간 방송은 청취하는데 있어 거부감을 나타낼 뿐 아니라, 멘트의 내용도 명확하게 들리지 않아 오히려 방해가 된다. 따라서 적당한 울림효과를 TTS 로 제작된 멘트에 적용하여 청취자들이 멘트에 집중하여 멘트의 내용을 파악할 수 있도록 한다.

울림효과는 사운드 효과 중에서 잔향효과를 적용하여 제작한다. 잔향효과의 잔향시간은 음의 울림의 양에 대한 가장 중요한 평가지수이며, 정상상태의 음이 60dB 감쇠하는 데까지 소요되는 시간으로 정의된다[8].

잔향효과 구현에 많이 사용되는 표준 라이브러리는 I3DL2 (Interactive 3D Audio, Level 2)이며, 잔향효과에서 사용되는 용어는 다음과 같다.

- 1. Direct Path
  - 적용된 효과가 없는 직접 음.
- 2. Early Reflection
  - 벽에 반사된 음에서 가장 빠른 음.
- 3. Late Reverberation
  - 벽에 반사되어 뒤늦게 들려오는 음.

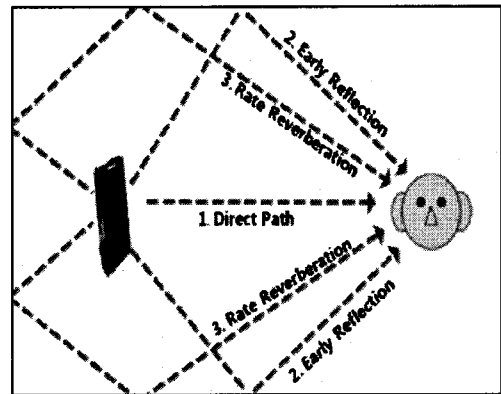


그림 2. 직접음과 반사음의 이동경로  
Figure 2. Moving path of directing sounds and reflecting sounds

그림 2는 방 안에서 직접 들리는 음과 반사되는 음의 이동 경로를 보여준다. I3DL2 에서 사용되는 용어들의 이동 경로를 보여준다.

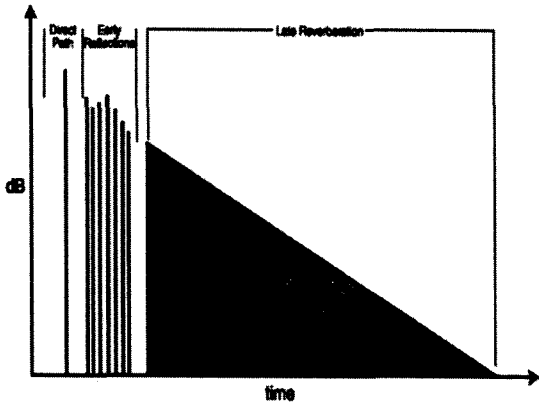


그림 3. 잔향효과의 임펄스 응답  
Figure 3. Impulse Response of Reverberation Effect

그림 3 은 원본 소리가 잔향효과필터를 통과하여 생긴 임펄스 응답을 보여준다. 직접 들려오는 음과 반사되어 들려오는 음들에 대한 데시벨 특징을 나타낸다.

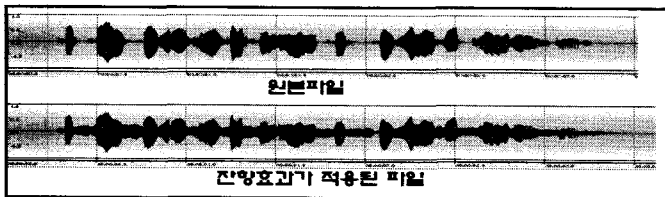


그림 4. 원본 파일과 잔향효과가 적용된 파일  
Figure 4. The original file and applicated reverberation effect file (Delay time = 0.5s)

그림 4 는 원본과 잔향효과를 적용한 경우의 파형을 보여준다.(잔향효과의 Delay 가 0.5 초로 적용한 파일이다. 총 시간이 0.5 초간 지연된 것을 볼 수 있다.)

가장 빠른 반사음(Early Reflection)과 벽에 반사되어 뒤늦게 들려오는 음(Late Reverberation)의 지연시간과 강도에 따라 잔향효과의 형태는 달라진다. 따라서 각 음의 지연시간과 강도를 각 매장에 맞추어 멘트를 변경시켜야 한다. 그러나 이러한 내용들을 정확하게 설정하기 위해서는 건축음향에 관련된 전문가와 전문장비로 모든 내용들을 측정하고 설계해야 한다. 그러므로 음향 설계에 관련이 없는 비전문가들도 이에 대해 평가할 수 있는 평가 지표가 필요하다.

#### 4. 음향 효과 평가 실험

음향 성능평가는 24 평 크기의 실습실을 이용하여 9 명의 20 대 남녀 학생들에 의해서 듣기 평가를 실시하였으며, 음향 성능은 “가청화를 이용한 소규모 다목적 홀의 음향 성능평가에 관한 연구”에서 소개된 평가도구들을 이용하여 실시하였는데 사용된 음향성능 구성요소는 표 1 과 같다[9].

표 1. 대상 다목적 홀의 음향성능 구성요소  
Table 1. Sounds performance component of the Multi-purpose hall

사용된 어휘	주관적 요소	물리적 요소
건조하다. / 충만하다.	음의 울림	잔향시간
먼곳에서 연주하는 느낌이다. / 가까운곳에서 연주하는 느낌이다.	음의 친밀감	초기 지연 시간
탁하다. / 명료하다.	음의 명료성	명료도
둔하다. / 선명하다.	음의 선명함	
날카롭다. / 부드럽다.	음의 균형	저음비
음이 한쪽으로 치우친다. / 음이 균형이 있다.	음의 균형	공간감
한쪽으로 집중된다. / 넓게 확산된다.	음의 확산감	

제안된 평가어휘를 바탕으로 미국의 심리학자 Osgood 에 의해 제안된 의미분별법(SD: Method of Semantic Differential)을 이용한 청취 실험에 사용될 설문지를 표 2 와 같이 구성한다.

표 2. 주관적 반응을 평가하기 위한 평가시트<sup>5)</sup>  
Table 2. Estimated sheet for subjective response estimate

항목별	평가	
	전혀그렇지않다 ←	→ 매우그렇다.
음의 울림(잔향감)	1 — 2 — 3 — 4 — 5 — 6 — 7	
음의 친밀감	1 — 2 — 3 — 4 — 5 — 6 — 7	
음의 명료성	1 — 2 — 3 — 4 — 5 — 6 — 7	
음의 선명함	1 — 2 — 3 — 4 — 5 — 6 — 7	
음의 포근함	1 — 2 — 3 — 4 — 5 — 6 — 7	
음의 균형	1 — 2 — 3 — 4 — 5 — 6 — 7	
음의 확산감	1 — 2 — 3 — 4 — 5 — 6 — 7	

실험은 두 개의 스피커가 설치된 실험실(7.8Mx9.5M) 에서 실시하였다. 표 1 을 기준으로 청취자는 그림 5 와 같은 위치에서 청취 성능을 판단하여 평가시트에 설문하였다.

5) 평가도구에는 음의 크기 부분이 있으나 볼륨 조절로 방송이 가능하여 이 부분은 제외하였다.

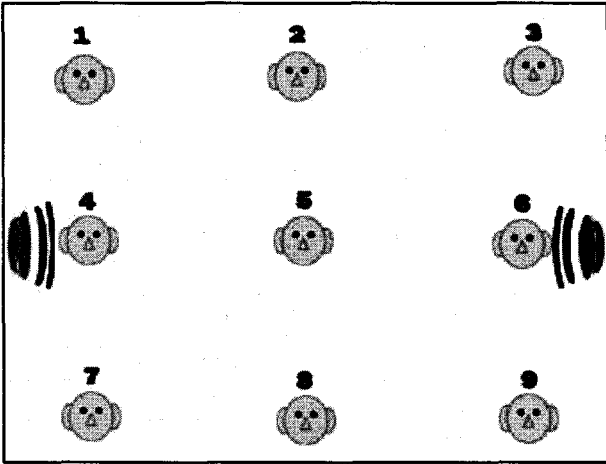


그림 5. 실험 공간에서의 청취자 위치  
Figure 5. Location of listeners in the testing space

5. 실험 결과 및 평가

본 논문에서 제안한 음향 효과는 Visual Studio 2008, C#, DirectX 를 이용하여 구현하였으며, 운영체제는 WindowsXP 이다.

동일 조건에서 잔향효과가 없는 경우와 있는 경우를 여러 번 시험하여 평균 결과 값을 비교하였다. 표 3 은 잔향효과가 없는 경우 실험결과이다.

표 3 잔향효과가 없는 경우 청취 평가 결과(평균치)  
Table 3. Hearing test result without reverberation effect (average)

위치(청취자)	1	2	3	4	5	6	7	8	9
음의울림(잔향감)	1	1	1	1	1	1	1	1	1
음의 친밀감	2	2	2	2	4	2	2	2	2
음의 명료성	7	7	7	7	6	7	7	6	7
음의 선명함	7	6	7	6	7	7	7	7	7
음의 포근함	2	2	2	1	3	1	2	3	2
음의 균형	1	5	2	1	6	2	3	6	2
음의 확산감	1	1	1	1	1	1	1	1	1

표 3 의 내용을 보면 전반적으로 명료성과 선명함은 높게 나왔다. 그러나 나머지의 평가 지표에 대한 결과는 대체적으로 낮은 편이며, 특별히 음의 균형의 경우는 몇 위치를 제외하고는 매우 낮게 나타났다.

잔향효과에서 잔향의 크기(dB)와 잔향시간을 조절하여 시험을 실시하였다. 잔향의 크기가 너무 높은 경우 사람들이 대체적으로 친밀감에서 매우 낮은 점수를 주었으며, 거부감이 매우 높았다. 그리고 명료성과 선명함의 수치가 매우 떨어져 잔향의 크기를 적당히 높여야 좋은 효과를 나타낼 수 있었다. 잔향시간이 길어지면 음의 균형에서 수치가 높아지는 결과를 가져왔으나 너무 길어질 경우 친밀감이 떨어지는 현상이 발생하였다. 표 4 는 사람들이 가장 적절하게 느끼는 잔향효과에 대한 청취 평가 결과표이다.

그림 6 은 원음과 잔향효과를 적용한 경우에 대한 실험 결과인 표 3 과 표 4 의 내용을 비교한 것이다.

그림 6 을 보면 전체적으로 음향 효과를 적용한 경우가 수치가 높다. 이는 청취자들이 음향 효과가 적용된 소리에 더 좋은 반응을 나타내는 것을 알 수 있다. 음의 울림과 음의 확산의 수치는 높으면 거부감을 가지게 된다. 이 두 가지 요소는 잔향효과를 많이 주면 거부감이 높고 음의 울림과 확산의 수치도 높아졌다. 따라서 잔향효과와 적당한 수치는 음의 울림과 확산의 수치가 중간에 나타날 때 사람들이 가장 좋아 하는 소리로 느껴진다는 결과를 얻을 수 있었다.

표 4 반응이 가장 좋은 잔향효과 적용 시 청취 평가 결과  
Table 4. Hearing test result using best reverberation effect (average)

위치(청취자)	1	2	3	4	5	6	7	8	9
음의울림(잔향감)	3	3	5	2	4	3	4	3	3
음의 친밀감	7	5	7	7	6	6	6	7	7
음의 명료성	6	5	5	5	5	5	6	6	7
음의 선명함	6	5	5	6	6	5	7	7	6
음의 포근함	6	5	4	7	5	4	5	6	6
음의 균형	7	5	6	7	6	6	7	7	7
음의 확산감	3	5	4	1	4	3	3	3	4

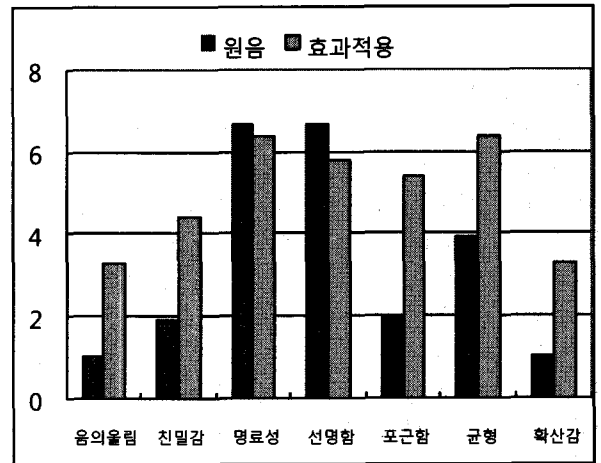


그림 6. 원음과 음향 효과를 적용한 경우 청취 평가 결과 비교

Figure 6. Comparing hearing test result between original sounds and applied sound effect sounds

실험 결과 명료성과 선명함에 있어서는 효과를 적용한 소리가 원음보다 떨어졌다. 그러나 실험에 참가한 청취자들은 멘트의 내용을 인지하는 데는 별 차이가 없었으며, 잔향음이 들어간 소리를 원음 소리보다 선호한다는 평가를 하였다.

이러한 결과를 토대로 잔향효과가 적용된 소리에 좀 더 집중하는지에 대한 실험을 실시하였다. 실험 내용은 청취자들이 자연스럽게 대화를 나누고 대화에 집중하게 되었을 때 청

취자가 느낄 수 있는 크기로 원음 멘트와 잔향효과가 적용된 멘트를 한번씩 번갈아 가면서 방송하는 것으로 하였다.

소리가 일정 크기로 높아졌을 때는 두개의 소리에 모두 인지하였으나 일정 크기보다 작은 경우에는 잔향효과가 적용된 소리에만 방송되는 것을 인지하였다.

이러한 결과를 토대로 멘트에 잔향효과가 적용되었을 경우 실질적으로 매장음원방송 서비스에서 이벤트나 사람을 찾는 안내방송등의 중요한 멘트를 사람들에게 전달할 경우 효과적으로 제공할 수 있음을 알 수 있었다.

## 6. 결론

본 논문은 TTS 로 만들어진 멘트를 각 매장에서 방송할 때, TTS 는 변경하지 않고 각 매장 특성에 따라서 멘트 내용을 잘 인지할 수 있게 하기 위한 방법에 관한 것이다.

사람이 소리에 대한 주관적인 인지도가 잔향효과에 많은 영향을 미친다는 건축음향관련 논문을 보고 이를 이용하여 TTS 로 만들어진 멘트 파일에 잔향효과를 적용하여 실험을 하였다. 실험결과 사람들이 잔향효과가 많이 적용된 경우 소리에 대한 거부감을 많이 나타내었고, 적당한 잔향 수치를 적용한 경우 듣기 좋다는 평을 내렸다. 이에 얻은 수치를 이용하여 적절한 잔향효과를 적용한 경우 각테일 파티효과가 나타남을 확인할 수 있었다. 즉, 적절한 잔향효과가 들어간 소리에 인지도가 높다는 결과를 얻었다.

상기 실험 결과를 토대로 매장에서 중요한 내용의 멘트 방송을 할 때 매장에 적절한 잔향효과를 삽입하여 방송하면, 고객들이 방송에 집중하게 되어 멘트 내용이 잘 전달된다는 결론을 얻었다.

## 감사의 글

이 논문은 중소기업 산학협력실 지원사업의 연구 결과로 수행되었습니다.

## 참 고 문 헌

- [1] Chang, M. S. Kang, S. M. (2007). "Implementation of Music Broadcasting Service System in the Shopping Center Using Text-To-Speech Technology" *Speech Sciences* Vol. 15, No. 4 (장문수, 강선미. (2007). "TTS 를 이용한 매장 음악 방송 서비스 시스템 구현.", *음성과학* 제 14 권 제 4 호.)
- [2] Kang, S. M. Kim, H. D. Chang, M. S. (2008) "Improvement of Shop Musix Broadcasting Services Using Music Lists and User Experience", *Speech Sciences* Vol.15, No.4 (강선미, 김현득, 장문수. (2008). "방송목록과 사용자경험정보를 이용한 매장음원방송 서비스의 개선.", *음성과학회*. 제 15 권 제 4 호)
- [3] Oh, Y. H. (1998). *Speech Language Information Processing*, Hongrung Publishing Company (오영환. (1998). *음성언어정보처리*. 흥릉과학출판사.)
- [4] Hyun, T. S. Choi, M. H. Kim, N. D. (2009) "An Architectural Planning of Auditoriums with Regard to the

Control for Echo, Sound Focus and Reverberation Time", *The Regional Association of Architectural Institute of Korea*. Vol.11 No.1 Serial Number 37

- (현택수, 최무현, 김남돈. (2009). "반향 및 음집중현상과 잔향시간조절을 고려한 공연장 건축 계획.", *대한건축학회지회연합회논문집*. 제 11 권 제 1 호 통권 37 호)
- [5] Lee, C. W. Kim, Y. H. Seo, C. K. Jeon, J. Y. (2008), "The Effects of Stage Reflectors on Stage Acoustics in Concert Halls" *J.Korean. Soc. Living. Enviro. Sys.* Vol. 15 No. 3 (이창우, 김용희, 서춘기, 전진용. (2008.) "콘서트 홀 무대음향에 대한 음향 반사판의 영향.", *한국생활환경학회지*, 제 15 권 제 3 호)
- [6] Ju, D. H. Yun, J. H. Kim, J. S. (2001) "Architectural Acoustics Design of H. Event Hall", *The Regional Association of Architectural Institute of Korea*, Vol. 21 No. 1 (주덕훈, 윤재현, 김재수. (2001). "H 다목적홀의 건축음향 설계." *대한건축학회 학술발표대회 논문집* 제 21 권 제 1 호)
- [7] Kang, S. H. (2003), "Increasing speech articulation by the loudspeaker directivity control", *Acoustical Society of Korea*, Vol. 22 No. 2 (강성훈. (2003.). "스피커의 지향성 제어에 의한 명료도 향상." *한국음향학회 학술발표대회 논문집*, 제 22 권 제 2 호)
- [8] Choi, D. Kim, D. G. Kim, J. S. (2008), "Evaluation on Architectural Acoustic Performance of J Art Center for Improvement of Acoustic Performance", *Korean Society of Environ. Engineers* (최 돌, 김대군, 김재수. (2008.) "음향성능개선을 위한 J 예술회관의 건축음향성능평가." *대한환경공학회 학술발표대회*.)
- [9] Ju, D. H. Yun, J. H. Kim, J. S. (2007.) "A Study about Appraise on Acoustic Performance of small-scaled Multi-Purpose Hall, using Auralization", *The Korean Society for Noise and Vibration Engineering (KSNVE)* Vol. 2007 No. 0 (주덕훈, 윤재현, 김재수. (2007.) "가청화를 이용한 소규모 다목적 홀의 음향성능평가에 관한 연구." *한국 소음진동 공학회 학술발표대회*. Vol. 2007 No. 0)

### • 강선미 (Kang, Sunmee)

서경대학교 전자공학과  
서울시 성북구 서경로 124(정릉동 16-1) 북악관 602호  
Tel: 02-940-7291 Fax: 02-919-5075  
Email: smkang@skuniv.ac.kr

### • 김현득(Kim, Hyundeuc)

서경대학교 전자컴퓨터공학과  
서울시 성북구 서경로 124(정릉동 16-1) 북악관 602호  
Tel: 02-940-7291 Fax: 02-919-5075  
Email: deempl@nate.com

### • 장문수 (Chang, Moonsoo) 교신저자

서경대학교 소프트웨어학과  
서울시 성북구 서경로 124(정릉동 16-1) 한림관 914호  
Tel: 02-940-7509 Fax: 02-919-5075  
Email: cosmos@skuniv.ac.kr