

음악 연주 휴머노이드 로봇

I. 서론

1920년 극작가 Karel Čapek가 발표한 ‘Rossum’s Universal Robots’에서 ‘Robot’이란 단어가 처음 사용된 이후로 로봇에 대한 개념과 정의가 공론화 되면서 많은 연구자들에 의해 개발되고 연구되어 왔다. 그 후 최초의 산업용 로봇이 개발 되면서 20세기까지 로봇은 단순하게 제조공장에서 조립, 용접, 핸들링 등을 수행하는 산업용 로봇을 지칭하는 말이었다. 그러나 21세기에 들어서면서 로봇 산업의 급속한 발전과 기능의 다양화로 각 역할에 맞게 분류되었으며, 특히 최근 들어서 서비스 로봇에 대한 연구가 각광받고 있다. 이러한 서비스 로봇 분야에서는 사람-로봇 상호작용(Human-Robot Interaction, HRI)에 대한 연구가 집중되고 있으며, 특히 감정표현, 인간 친화적인 상호 정보 전달을 할 수 있는 다양한 엔터테인먼트 로봇에 대한 산업 및 연구가 주목받고 있다. 이러한 로봇들은 사람의 의지를 인식하기 위하여 시각, 청각, 촉각 등을 감지할 수 있는 센서를 사용한다. 이렇게 엔터테인먼트 로봇 분야에서 시각 및 촉각 센서는 급속한 발전을 이루어 사람과의 상호 작용이 가능한 정도로 기술이 발전했지만, 아직 청각을 통한 사람의 의지를 인식하기 위한 센서 및 알고리즘은 미약한 상태이며, 때문에 현재 사람의 음성과 다양한 소리들을 인식하는 연구들이 활발히 진행되고 있다^[1]. 이러한 소리를 인식하는 연구 중에서 auditory feedback 연구는 의료 분야에서만 국한적으로 사용되고 있다^[2]. 특히, 말을 더듬는 사람이 본인의 소리를 듣고 상태를 이해하여 더듬는 증상을 자가 치료하는 방법이 있다^[3].



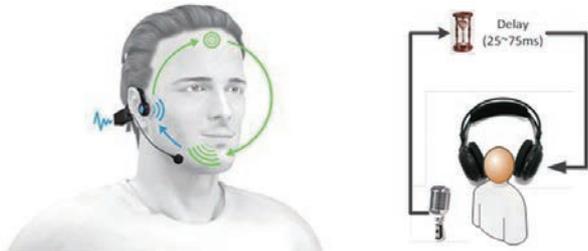
박 현 준
경희대학교
전자·전파공학과 석사과정



김 동 한
경희대학교 전자정보대학
전자·전파공학과

1. 관련 선행 연구

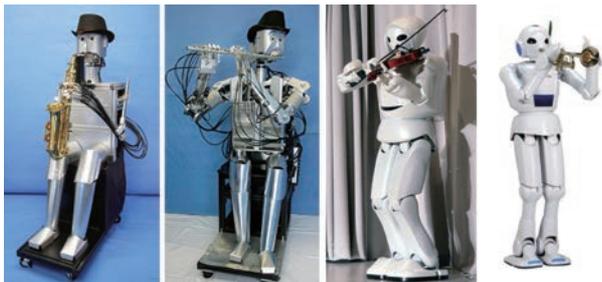
Waseda 대학에서는 ‘Waseda Saxophonist Robot’^[4]과 ‘Flutist



〈그림 1〉 Auditory feedback의 원리

Robot^[5]을 개발했다. 이 로봇들은 공기 펌프 등을 이용하여 인위적인 공기의 흐름 제어를 통해 이를 관악기에 흡기시켜 연주한다. 이는 사람의 장기인 폐를 흉내냈다는 것에서 큰 의미를 찾을 수 있다. 하지만, 주로 사람의 연주와 흡사한 기계적인 장치만을 모사했다는 한계가 있다. 다른 종류의 악기 연주로봇으로써 Toyota에서는 정밀한 제어 기술을 이용한 바이올린 로봇을 개발하여, 실제 시연을 하였다^[6]. Ryukoku 대학에서는 ‘Kansei’를 이용하여 바이올린 연주로봇에 사람의 감정을 반영한 제어 명령을 넣어 좀 더 자연스러운 바이올린 연주를 하도록 연구했다^[7-9]. 하지만, 이러한 연구에도 로봇이 연주한 소리를 로봇이 제어 명령에 피드백하지 않기 때문에 정해진 명령을 따라만 가는 단순한 HRI를 가지고 있다. 이러한 로봇들의 경우는 로봇 하드웨어 설계와 제어기술의 우수성을 보여주는 사례로 충분한 가치가 있지만, pre-programmed 된 제어 명령을 정확하게 따라가기만 한 것이다.

따라서 이러한 한계점을 개선하기 위해 본 연구에서는 앞서 소개한 auditory feedback 기술을 HRI의 중요 요소라고 판단하였다. 따라서 기존에 있던 바이올린 연주



〈그림 2〉 선행 연구된 음악연주 로봇

로봇을 개선하고, auditory feedback 시스템을 접목시킨 음질 평가 시스템에 대해 소개하려고 한다. 2장에서는 바이올린 로봇 시스템에 대해서 소개하고, 3장에서는 auditory feedback을 접목한 음질 평가 시스템에 대해서 소개한다. 마지막으로 4장에서 결론 및 연구방향으로 끝을 맺는다.

II. 바이올린 로봇 시스템

1. 바이올린 로봇의 소개

본 연구에서 개발된 로봇은 6축 다관절 산업용 로봇팔(RV-2SD)을 이용하여 제작하였다. 제작된 바이올린 로봇은 〈그림 3〉과 같으며, 사람과 유사한 형태를 취하고 있다. 바이올린은 성인들이 사용하는 4/4 크기를 사용하며 외부 잡음에 강한 전자 바이올린을 사용한다. 고정을 위해 바이올린의 형태를 변형할 경우 악기의 고유 소리가 변하기 때문에 최대한 바이올린에 손상을 주지 않는 흡착식 방식으로 고정했다. 로봇의 오른팔의 end-effector는 현악기를 연주할 때, 활을 잡는 역할을 한다. 이때 활과 현이 적절한 압력으로 눌러지도록 지탱할 수 있게 설계를 하였다. 활의 위치에 따라 소리의 음질이 바뀔수 있기 때문에 현악기 연주 시 음의 보정을 위하여 활의 위치를 조절할 수 있는 장치를 설계하였다.

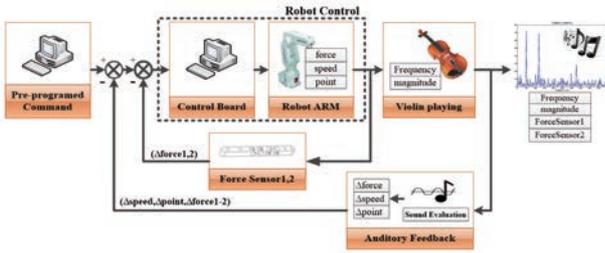
Auditory feedback 기술은 연주 로봇에게 있어서 HRI의 중요요소이다.

2. Auditory feedback 시스템의 구조

앞서 언급된 바와 같이 auditory feedback은 의료 분야에서 말더듬이 치료 방법으로 주로 연구되었다. 이리



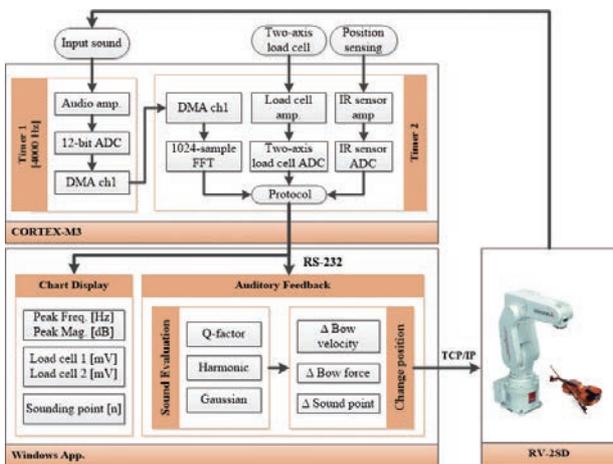
〈그림 3〉 제작된 바이올린 연주 로봇



〈그림 4〉 Auditory feedback 시스템의 흐름도

한 치료 방법은 환자의 목소리를 마이크로 녹음하고, 25~74ms 지연시켜 헤드폰을 통해 다시 듣는 과정을 통하여 말더듬이 증상을 치료한다. 이러한 기법을 바이올린 로봇에 적용하여 음질을 스스로 향상시킬 수 있는 방법을 제안한다. 〈그림 4〉는 이러한 바이올린 연주 로봇의 시스템 흐름도를 나타낸다.

왼쪽의 첫 블록에서는 바이올린 연주와 관련된 프로그래밍 명령을 제어보드로 실시간으로 전달한다. 이와 동시에 auditory feedback 블록에서는 바이올린 연주 소리를 입력 받아 A/D(Analog/Digital) 변환과 STFT(Short Time Fourier Transform)를 거쳐 소리의 질을 판단한다. 이를 바탕으로 로봇으로 들어가는 입력의 변화 값을 결정하여 미리 프로그램된 명령어에서 빼게 된다. 이와 같은 방식으로 바이올린 연주의 소리를 피드백함으로써 로봇은 단순히 프로그램된 명령에 따라서 “움직이는 것”이 아니라 “연주”를 하게 된다.



〈그림 5〉 바이올린 연주 로봇의 제어 구조

3. Auditory feedback 시스템의 구조

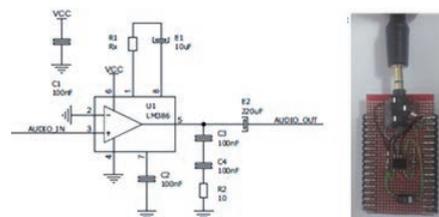
〈그림 5〉는 본 연구에서 사용하는 전체 제어 구조이다. 우선 마이크로프로세서 모듈로 실시간으로 현재 연주정보를 수집하고, 차트 도구상자에서 이를 표시한다. PC 응용 프로그램은 auditory feedback 시스템이 있으며, 이로 인해 현재 소리를 판단하여 바이올린 연주 로봇의 움직임을 제어한다. 마이크로프로세서 모듈은 ARM cortex-M3를 사용하여, 실시간으로 활의 힘, 활의 속도, 연주점 등을 입력 받을 수 있다. 이 모듈은 오디오 지각 시스템, 활의 힘 측정 시스템, 연주점 측정 시스템과 연동되어 있다. 오디오 지각 시스템에서는 실시간으로 연주되는 소리를 STFT과정을 거쳐 기음주파수와 크기를 얻는다. 활의 힘 측정 시스템에서는 연주하는 동안의 활의 누르는 힘을 얻고, 연주점 측정 시스템에서는 활의 연주되는 위치를 얻는다. 이 시스템들로 인해

얻는 값들을 하나의 프로토콜로 만든 후, PC 응용프로그램으로 전송한다.

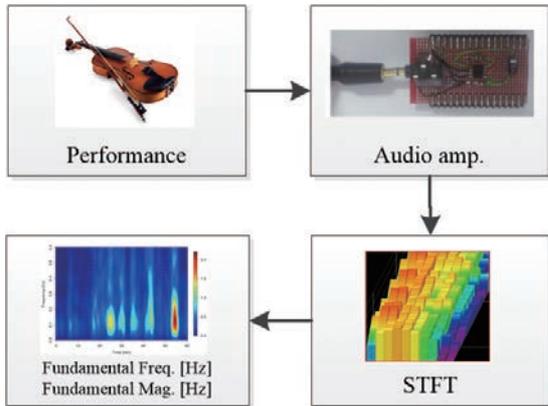
4. 오디오 지각 시스템

바이올린 소리에는 기본 주파수와 하모닉 주파수를 동시에 포함한다. 이때 연주음의 음계는 기본 주파수와 크기에 의해 결정된다. 따라서 실시간으로 소리를 듣고, 기본 주파수와 크기 값을 추출할 수 있는 오디오 지각 시스템을 개발한다. 연주되는 소리를 실시간으로 획득하기 위해 다음 〈그림 6〉과 같은 저전력 오디오 앰프를 이용한 지각 시스템을 개발한다.

앞서 언급한 바와 같이 본 연구에서는 STFT를 구현하여 시간영역에서의 정보와 주파수 영역에서의 정보를 동



〈그림 6〉 오디오 지각 시스템의 회로도 및 제작된 모습

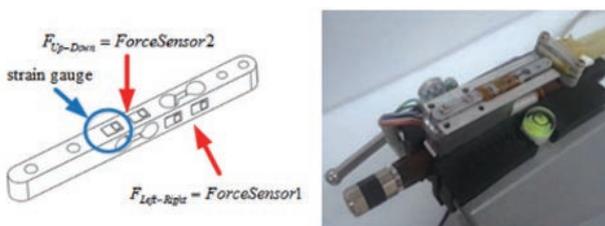


〈그림 7〉 오디오 지각 시스템의 흐름도

시에 살펴보는 것이 가능해진다. 이 해석기법은 소리를 빠르게 분석할 수 있고, 비교적 간섭에 대한 영향이 적은 장점이 있다. 〈그림 7〉과 같이 바이올린 소리는 픽업 장치를 사용하여 현의 진동을 아날로그 전기신호로 변환되어 출력되고 오디오전용 앰프로 소리를 증폭시켜 4000Hz 샘플링 주파수로 입력 받는다. 본 시스템에서는 30Hz마다 샘플링하여 FFT를 수행한다. 설정된 Timer를 사용하여 연주되는 소리를 12bit A/D 변환하고, DMA(Direct Memory Access, DMA)기능으로 원하는 샘플링 시간마다 FFT 처리하여 STFT 기능을 구현했다. 이와 같은 과정으로 기음 주파수와 크기를 구할 수 있다.

5. 활의 힘 측정 시스템

바이올린은 활과 브릿지 사이의 거리에 따라 활의 힘이 변하고 소리의 음질을 결정짓는 중요한 요소이다. 따라서 실시간으로 활의 힘을 측정할 필요가 있다. 이를 위해 〈그림 8〉과 같이 힘 센서를 제작했다.



〈그림 8〉 2축 힘 센서의 개념도 및 제작된 모습



〈그림 9〉 제작된 바이올린 연주 로봇

활의 누르는 힘과 비틀리는 힘을 측정하기 위해 로봇 팔 끝단에 활 손잡이를 제작하여 힘 센서를 장착하였다. 현재 제작된 로드셀은 쌍안경 구조가 2개가 존재한다. 이는 각각 활의 누르는 힘과 비틀리는 힘을 측정하며, 각 쌍안경 구조 중심축의 수직방향 평면에 부착한다. 로드셀에 외부 힘이 작용할 경우 로드셀이 변형되고 이로 인해 부착된 스트레인게이지 값이 변형된다. 이때 스트레인게이지의 저항에 전압을 인가하면, 출력단자에서 스트레인게이지의 저항값의 변화에 비례하는 전압이 출력된다. 이렇게 출력된 전압 값을 로드셀 전용 op-amp로 증폭한 후, 마이크로프로세서에 전달한다.

6. 활의 연주점 측정 시스템

바이올린 연주 시, 기음주파수는 현의 장력, 현의 길이, 선밀도에 따라 달라진다. 또한 연주 점은 브릿지에서 멀어질수록 기음주파수가 감소하며, 이로부터 연주하고자 하는 음마다 적합한 연주점이 결정된다. 본 연구에서는 이러한 연주 점을 파악하기 위하여 〈그림 9〉와 같이 10mm 간격으로 반사형 포토 인터럽트 5개를 일렬로 배치한다. 이 측정 장치로 인해 실시간으로 연주 점을 측정한다.

III. G.H.Q 음질 평가 시스템

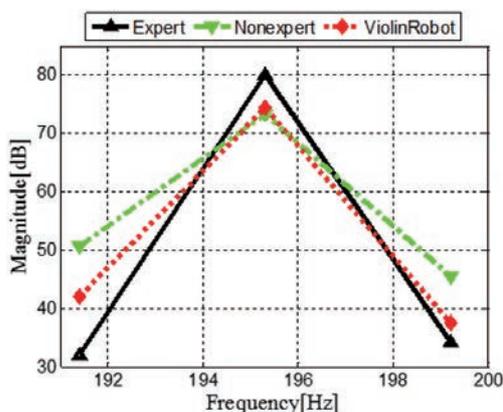
바이올린의 소리는 전문가들의 평가로 전체적인 소리를 평가를 받기 때문에 바이올린 음질을 주관적으로 평가된다. 본 논문에서는 바이올린의 음질을 정량적으로 평가하기 위해 G.H.Q 음질 평가 시스템을 개발했다.

G,H,Q(Gaussian-Harmonic-Quality) 음질 평가 시스템은 바이올린의 소리 평가를 객관적으로 평가하는 시스템이다. 바이올린 전문가들의 연주 결과와 바이올린의 특성을 고려해 만들어진 시스템이다. 바이올린의 소리는 기음과 배음이 복합된 복합음으로 구성되어 있으며 바이올린의 연주점, 활의 속도/힘에 따라 기음과 배음 주파수가 다르게 발생한다^[6]. 기음 주파수는 소리의 음계와 음정을 나타내고 배음 주파수는 소리의 음색을 결정한다. 이러한 요소를 이용하여 소리의 음질을 평가한다. 또한 본 시스템은 바이올리니스트와 바이올린 로봇에 적용해 다양한 연주점, 활의 속도, 활의 힘에 따라 소리를 평가해 시스템의 음질 평가에 대한 신뢰성을 확인했다.

우선 Gaussian 인자를 이용한 음질 평가에 대해 소개하겠다. Gaussian 인자는 연주되는 소리의 기음 주파수와 크기의 평균값으로 전반적으로 음을 일정하게 연주했는지에 대해 음질의 정확도를 평가한다.

모든 음원의 소리에는 기음과 배음이 존재한다. 음악에서의 악보에 표기된 음계는 기음 주파수를 의미하고 음정을 결정한다. 배음 주파수들은 기음의 음색을 결정짓고 음색을 풍부하게 해준다.

이러한 기음과 배음 주파수의 조합에 따라 음질이 달라지며, Harmonic 인자는 이러한 기음 주파수와 배음 주파수의 비율에 따라 음질을 평가한다. 이때 바이올린의 경우 기음 주파수와 배음 주파수의 이상적인 비율은 2이다.



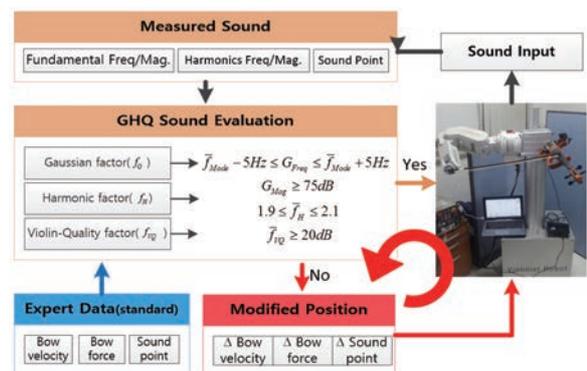
〈그림 10〉 연주자에 따른 Q-인자의 음질 평가 결과

Quality 인자는 회로나 시스템에서 공진의 주파수와 날카로움을 나타내는 인자를 응용해 음질을 평가한다. 본 연구에서는 이러한 Quality 인자를 응용해 연주되는 소리의 기음 주파수를 평가하는 Violin-Quality factor(f_{VQ})를 개발했다. f_{VQ} 는 연주하고자 하는 소리의 기음 주파수와 주변주파수의 크기들을 이용해 기음주파수의 크기로 대역폭 값을 구한다. 이 값은 음의 날카로움과 선명도를 나타내며 연주되는 소리의 기음주파수를 분석해 연주하는 음을 평가하는 지표로 활용된다. 〈그림 10〉은 바이올린 전문가, 바이올린 초보자 그리고 로봇의 연주를 Q-factor 평가 방법을 통해 분석한 결과이다. 바이올린 연주 실력에 따라 음의 날카로움과 선명도가 다르다는 것을 확인할 수 있고, Q-factor 평가에 대한 신뢰성을 검증했다.

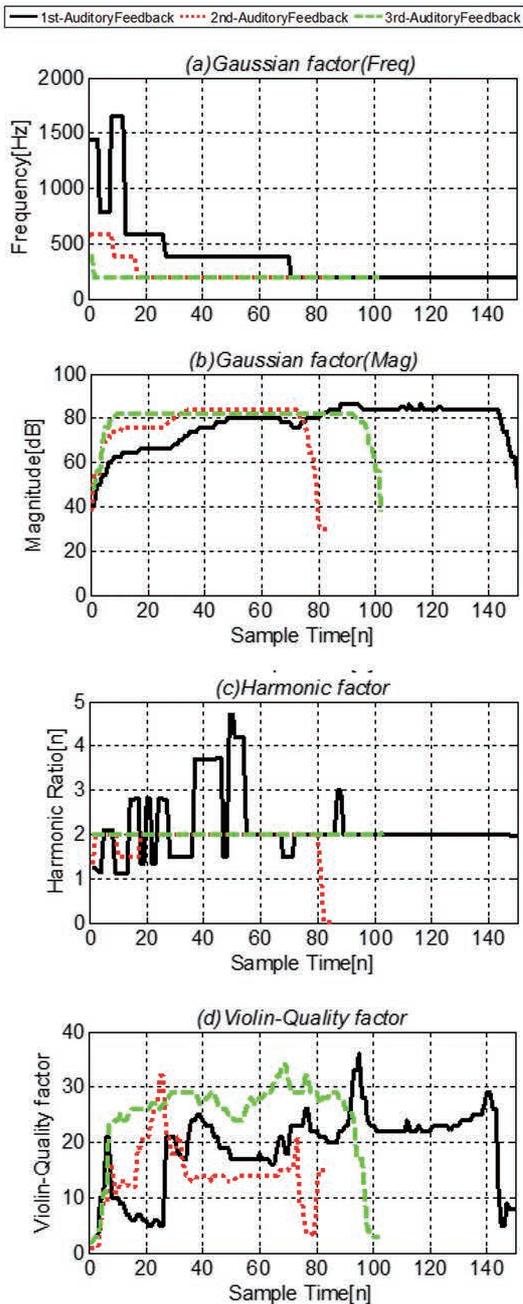
IV. 실험 및 결과

Auditory feedback 구조를 적용시킨
바이올린 연주 로봇의 연주음은
G.H.Q 음질 평가 시스템의 모든 조건에
만족하는 결과를 얻었다.

바이올린 연주자의 행동을 모방한 auditory feedback 구조를 적용시킨 바이올린 로봇 시스템의 흐름도는 〈그림 11〉와 같다. 실제 사람 연주자의 청각 피드백 모방한 제어시스템과 음질을 판단할 수 있는 G,H,Q 음질 평가 시스템을 바이올린 연주 로봇에 접목시켜 초기 음을 듣고 음질을 판단하여 스스로 연주점의 위치, 활의 힘, 활의 속도를 추정하여 다음 연주 시 좋은



〈그림 11〉 Auditory feedback 시스템의 흐름도



〈그림 12〉 Auditory feedback 적용 실험 결과

음이 나도록 연주점과 활의 힘, 속도를 변경한다. 이러한 과정을 반복하며 G,H,Q 조건에 만족한 경우, 연주된 소리가 좋다고 평가한다.

〈그림 12〉는 실제 바이올린 로봇에 auditory feedback 시스템을 적용시킨 결과이다. 1번째 실험에서 3번째 실험으로 갈수록 음질의 부족한 점을 스스로 판단하고, 바

이올린 전문가의 표준 정보와 비교하여 연주점과 활의 속도를 변경하는 결과를 보이며, 3번째 실험에서 G,H,Q 음질 평가 시스템의 모든 조건에 만족하는 결과를 얻었다.

V. 향후 연구 및 결론

지금까지 auditory feedback 시스템을 접목한 바이올린 연주 로봇에 대하여 소개하였다. 향후 사람과 협연이 가능한 로봇 시스템으로 발전시킬 예정이며, 많은 실험을 통해 데이터를 수집하여 G,H,Q 음질 평가 시스템의 신뢰도와 정확도를 더욱 향상시킬 것이다. 추가적으로 여러 명의 바이올린 전문가를 초빙하여 바이올린 로봇 연주 결과를 평가 받을 것이다.

참고 문헌

- [1] Fujita, M., & Kitano, H. (1998). Development of an autonomous quadruped robot for robot entertainment. *Autonomous Robots*, 5(1), 7–18.
- [2] Yates, Aubrey J. Delayed auditory feedback. *Psychological bulletin*, 1963, 60(3), 213–232.
- [3] Cai, S. Online control of articulation based on auditory feedback in normal Speech and stuttering: behavioral and modeling studies, 2011, Diss. Massachusetts Institute of Technology
- [4] Solis, J., Takanishi, A., & Hashimoto, K. (2010). Development of an Anthropomorphic Saxophone-Playing Robot. In *Brain, Body and Machine* (pp. 175–186). Springer Berlin Heidelberg.
- [5] Solis, J., Chida, K., Suefuji, K., & Takanishi, A. (2006). The development of the anthropomorphic flutist robot at Waseda University. *International Journal of Humanoid Robotics*, 3(02), 127–151.
- [6] Kusuda, Y. (2008). Toyota's violin-playing robot. *Industrial Robot: An International Journal*, 35(6), 504–506.
- [7] Shibuya, K., Matsuda, S., & Takahara, A. (2007, August). Toward developing a violin playing robot—bowing by anthropomorphic robot arm and sound analysis. In *Robot and Human interactive Communication*, 2007. RO-MAN 2007. The



16th IEEE International Symposium on (pp. 763–768). IEEE.

[8] Shibuya, K., Asada, T., & Sugano, S. (1998, October). An algorithm to convert KANSEI data into human motion. In Systems, Man, and Cybernetics, 1998. 1998 IEEE International Conference on (Vol. 2, pp. 1190–1194). IEEE.

[9] Shibuya, K., Matsuda, S., & Takahara, A. (2007, August). Toward developing a violin playing robot—bowing by anthropomorphic robot arm and sound analysis. In Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on (pp. 763–768). IEEE.

[10] Shibuya, K., Matsuda, S., & Takahara, A. (2007, August). Toward developing a violin playing robot—bowing by anthropomorphic robot arm and sound analysis. In Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on (pp. 763–768). IEEE.

[11] Wonse Jo, Hyeonjun Park, Bumjoo Lee, Donghan Kim. A study on improving sound quality of violin playing robot. In: Automation, Robotics and Applications (ICARA), 2015 6th International Conference on, IEEE, 2015, p. 185–191.



박현준

- 2014년 2월 호서대학교 학사(졸)
- 2014년 3월~현재 경희대학교 전자·전파공학과 석사과정

〈관심분야〉

Human-Robot Interaction, 생체모방로봇



김동한

- 1996년 2월 한국과학기술원 학사(졸)
- 1998년 2월 한국과학기술원 석사(졸)
- 2003년 2월 한국과학기술원 박사(졸)
- 2003년 3월~2004년 2월 일리노이 주립대학교 Post Doc.
- 2004년 3월~2007년 2월 삼성전자 무선사업부 책임연구원
- 2007년 3월~현재 경희대학교 전자정보대학 교수

〈관심분야〉

로보틱스, Human-Robot Interaction