

# 피치 변화음의 합성을 위한 도파관 모델

## Pitch-shifted sound synthesis using digital waveguide model

조상진, 강명수, 정의필

Sang-jin Cho, Myeong-su Kang, Ui-pil Chong

### 요약

디지털 도파관 모델은 파동 방정식의 일반해를 이용하여 진행파를 표현하고 이 진행파의 파동이동을 지연 라인으로 나타낸다. 일반적인 도파관 모델에서의 단일 지연은 샘플링 시간 간격을 의미하지만, 공간 기준 도파관 모델의 단일 지연은 샘플링된 공간의 거리를 의미한다. 이러한 차이점으로 인해 파동의 이동 거리를 직접적으로 조절할 수 있는 공간 기준 도파관 모델이 비브라토 음과 같이 피치가 변하는 음을 합성할 수 있다고 알려져 있다. 본 논문에서는 지연라인의 길이의 비로서 피치가 변하는 음을 합성할 수 있는 시간 기준 디지털 도파관 모델을 제안하고 기존의 공간 기준 도파관 모델과의 성능을 비교하였다.

### Abstract

In the digital waveguide theory, traveling waves are represented by general solution to the wave equation that is second-order linear partial differential equation. The movement of these waves can be implemented using only delay lines. An unit delay in the general digital waveguide describes a sampling time interval. However, in the space-based digital waveguide the unit delay implies the spatial sampling distance. In consideration of these differences between two models, it is known that the space-based digital waveguide model is adequate to synthesize pitch-shifted sounds such as vibrato because the propagation distance can be directly control. In this paper, the time-based digital waveguide model which also synthesizes pitch-shifted sounds is proposed and compared with space-based digital waveguide.

**Keywords :** Digital waveguide model, Sound synthesis, Pitch shifting.

## I. 서론

물리적 모델링은 컴퓨터 음악과 악기 음 합성 분야에서 가장 빠르게 발전하고 있는 분야 중 하나이다. 멀티미디어 응용이 개인 컴퓨터로 손쉽게 구현됨에 따라 악기의 물리적 특징을 기반으로 한 가상 악기의 소비자층이 형성되었고 사운드 이펙트와 음 합성 알고리즘이 처음으로 MPEG4 멀티미디어 표준으로 채택되기도 하였다[1]. 의심할 것도 없이 물리적 모델링은 향후 디지털 오디오 분야에서 중요한 역할을 할 것이라는 것은 쉽게 예상할 수 있다.

물리적 모델링은 악기의 음 발생 원리를 분석, 시뮬레이션을 통해 합성하는 방식이다. Karplus-Strong (KS) 알고리즘은 2탭 평균 필터와 노이즈 발생기를 이용하여 뜯는 현악기와 드럼의 음색을 가진 소리를 합성하였다[2]. 이후

Smith가 이를 확장시켜 일반화 하였고 이를 바탕으로 디지털 도파관(digital waveguide) 이론을 발표하였다[3]. 이 이론에서는 파동 방정식의 일반해를 디지털 영역으로 옮겨 파동의 움직임을 지연라인에서 표현하였다. 파동 방정식의 일반해는 샘플링을 거쳐 이산 영역으로 표현할 수 있는데, 샘플링의 기준을 시간과 공간에 대해 각각 적용할 수 있다. 일반적으로는 일정한 시간 간격으로 샘플링하는데, 이 경우 파동은 단일 지연을 샘플링 시간 간격으로 이동하므로 파동의 이동 속도는 일정하다. 그리고 지연 라인의 길이는 파동의 주파수와 관련 있으므로 고정된 지연 라인의 길이를 갖는 도파관 모델은 일정한 피치를 갖는 음만 합성할 수 있다. 이에 반해 공간 기준 도파관 모델은 샘플링 기준을 공간 변수로 설정하였기 때문에 단일 지연이 샘플링된 공간의 거리를 표현한다[4,5,6]. 이는 파동이 단일 지연을 지나는데 걸리는 시간이 일정하지 않을 수 있다는 것이며, 또한 피치가 변하는 음을 합성할 수 있다는 의미도 내포하고 있는 것이다. 이를 정리하면, 공간 기준 도파관 모델은 고정된 지연 라인에서 파동의 속도를 변화시킬 수 있으므로 피치가 변하는 음을 합성할 수 있고, 시간 기준 도파관 모델은 고정된 지연 라인에서 파동의 속도를 변화시킬 수 없

\*울산대학교

접수 일자 : 2009. 3. 29      수정 완료 : 2009. 4. 28

게재확정일자 : 2009. 4. 29

※ 이 논문은 2007년 울산대학교 교내학술연구비 지원으로 수행되었습니다.

으므로 고정된 피치 음만 합성할 수 있다는 것이다[6]. 이와 관련하여 [6]의 저자는 그의 논문에 다음과 같이 기술하였다.

“디지털 영역에서 파동의 속도가 1로 고정되어 있다는 것은 실제 악기의 물리적 현상에 비추어 보면 적절하지 않다. 이러한 제약 사항 때문에 시간 기준 디지털 도파관 모델에서는 음의 높이 변화를 수반하는 연주 기법을 표현할 수 없다.”

본 논문에서는 시간 기준 도파관 모델의 고정된 지연 라인에서 파동의 이동을 합성음의 피치에 맞게 업 샘플링 또는 다운 샘플링하여 비브라토와 같이 피치가 변하는 음을 합성할 수 있는 모델을 제안한다. 본 논문의 구성은 다음과 같다. 2장에서는 공간 기준 도파관 모델을 3장에서는 제안한 도파관 모델을 설명한다. 4장에서는 각 모델의 합성 결과를 비교하고 5장에서는 결론과 향후 과제에 대해 기술한다.

## II. 공간 기준 디지털 도파관 모델

공간 변수를 기준으로 한 현악기의 디지털 도파관 모델에서 지연은 현의 물리적 거리를 표현한다. 공간 변수의 샘플링 간격을 나타내는 상수를  $X$ , 샘플링 시간 간격을  $\tau$ 라고 하면  $X = k\tau$ 의 관계식이 성립하며, 여기서  $k$ 는 공간상의 거리  $X$ 를 이동하는 파동의 속도이다. 이를 이용하여 파동 방정식의 일반해를 이산 영역으로 유도하면 식 (1)과 같다.

$$y(x_m, t_{nk/c}) = y_r(m-n) + y_l(m+n) \quad (1)$$

식 (1)에서 한 번의 신호 지연은 공간 축에서 신호가  $X$  거리만큼 이동함을 의미하고, 이 때  $k/c$ 의 시간이 소요됨을 의미한다. 파동의 속도와 이동 거리는  $X = cT$ 의 관계식이 성립하므로  $nk/c$ 를 다시 풀어쓰면 식 (2)가 성립한다.

$$nk/c = nT/\tau \quad (2)$$

식 (2)는 파동이 거리  $X$ 를 이동하는데  $T/\tau$ 의 시간이 걸리며,  $T/\tau = k/c$ 이므로 파동의 움직임은 파동의 속도 비율에 의해 표현된다고 할 수 있다. 따라서 식 (1)은 식 (3)과 같이 수정할 수 있다[6].

$$y(x_m, t_n T/\tau) = y_r(m-n) + y_l(m+n) \quad (3)$$

식 (3)에서 각 지연마다 발생하는 시간 축의 지연은  $T/\tau = k/c$ 이고, 디지털 도파관 모델에서  $T$ 는 최소 시간 단위로서 1이고, 파동의 속도  $c$ 도 항상 1이므로 상대 속도인  $k$ 값의 변화에 따라서 합성음의 주파수를 변화시킬 수 있다. 상대 속도  $k$ 는 식(4)와 같이 정의한다[4,6].

$$k = \frac{\text{기준 지연수}}{\text{합성음의 지연수}} = \frac{\text{합성음의 주파수}}{\text{기준 주파수}} \quad (4)$$

공간 기준 디지털 도파관 모델에서는 위에서 보인 것처럼 특정 음을 합성하기 위한 두 가지 방법이 있다. 하나는 지연 수를 고정하고 상대 속도를 변화시키는 것, 다른 하나는 상대 속도를 고정시키고 지연의 수를 변화시키는 것이다. 상대 속도를 고정시키는 방법은 시간 기준 디지털 도파

관과 동일한 방법으로 음을 합성하는 것이 된다[6]. 예를 들어, 기준 음을 G3(196Hz, 지연수 225)로 했을 때, G4음(391.99Hz, 지연수 112.5)을 합성하기 위해서는  $k = 225/112.5 = 391.99\text{Hz}/196\text{Hz} = 2$ 가 된다. 이러한 관계를 그림 1에 나타내었다.

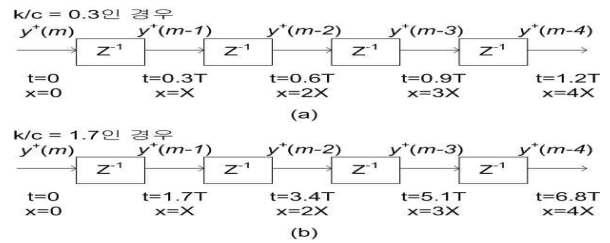


그림 1. 상대 속도와 샘플링 시간의 관계

Fig. 1. Relationship between relative wave speed and sampling time.

공간 기준 디지털 도파관 모델이 시간 기준 디지털 도파관 모델보다 파동의 움직임을 더 정확히 묘사할 수 있다. 극단적인 예로, 샘플링 주파수를 44100Hz로 했을 때, 피치가 44100Hz인 현악기 음을 시간 기준 디지털 도파관 모델을 이용한다면 지연의 수는 1이다. 이는 원하는 피치는 묘사할 수 있지만 단 하나의 지연이 현을 정확하게 묘사할 수 없다. 하지만 공간 기준 디지털 도파관 모델은 합성음의 주파수와 무관하게 지연의 수를 결정할 수 있는 장점이 있다.

## III. 제안한 시간 기준 디지털 도파관 모델

Smith가 제안한 1차 선형 양방향 디지털 도파관에서 이산 시간 상수  $n$ , 위치  $k$ 에서 도파관 출력은 식 (5)와 같다.

$$y(n, k) = y_r(n-k) + y_l(n+k) \quad (5)$$

여기서  $y_r(n)$ 과  $y_l(n)$ 은 각각 오른쪽과 왼쪽으로 이동하는 파동이다[7].

파동의 전파 속도는  $c = X/T$  로서, 공간과 시간 샘플링 간격의 비로 나타낼 수 있다. 전파 속도가 공간적으로 균일하게 분포한다고 가정하면 파동은 시간 상수  $n$ 과  $n+1$ 사이 에 거리  $\tilde{X} = c(n)T$  를 이동한다.  $c(n)$ 은 시간에 따른 파동의 속도의 변화를 의미하며, 만약 일정한 속도  $c$ 를 갖는다면  $c = c(n) = X/T$ 으로 이동 거리  $\tilde{X} = X$ 이다. 그림 2는 파동의 이동 거리를 나타내는 것으로  $c(n) > c$ 의 경우와  $c(n) < c$ 의 경우를 나타낸다.

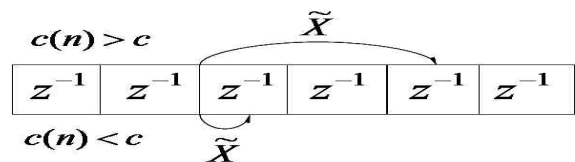


그림 2. 전파 속도에 따른 파동의 이동거리

Fig. 2. Propagation distance corresponding to wave speed.

디지털 도파관 모델에서 소수점 이동은 미소지연 필터를 이용하여 구현한다[8]. 간단한 설명을 위해 오른쪽으로 이동하는 파동,  $y(n,m) = y_r(n,m)$  만을 고려해 본다. 왼쪽으로 이동하는 파동에도 동일하게 적용할 수 있으며 손실이 없다고 가정한다. 초기 분포가  $y(0,m), m = -\infty, \dots, \infty$  으로 주어졌을 때,  $n=1$  이고 관측 지점이  $k$  인 곳에서의 출력은  $y(1,k) = y(0, k - c(0)T)$  이고,  $n=2$  인 경우  $y(2,k) = y(0, k - c(0)T - c(1)T)$  이다. 이를 일반화 하면 식 (6)과 같다.

$$y(n,k) = y(0, k - T \sum_{l=1}^{n-1} c(l)), n > 0 \quad (6)$$

그림 2에서와 같이 파동의 전파 속도가 일정하지 않아서  $c(n) = c + c_d(n)$  이라면, 식 (6)을 식 (7)과 같이 바꿔 쓸 수 있다.

$$y(n,k) = y(0, k - n - T \sum_{l=1}^{n-1} c_d(l)), n > 0 \quad (7)$$

여기서  $c_d(n)$  은  $0 \leq c_d(n) < c$  인 경우이며, 후항의 합계 항은 식 (8)과 같이 정의하여 미소지연 필터로 구현한다[9].

$$d(n) = - T \sum_{l=1}^{n-1} c_d(l) \quad (8)$$

그러나  $c_d(n) \geq c$  인 경우에는  $c_d(n)$  을 식 (9)와 같이 수정하여야 한다.

$$c_d(n) = i(n)c + c_f(n) \quad (9)$$

여기서  $i(n)$  은 양의 정수이고,  $0 \leq c_f(n) < c$  이다. 이 경우 식 (7)은 다음과 같이 수정된다.

$$y(n,k) = y(0, k - n - M - T \sum_{l=1}^{n-1} c_f(l)), n > 0 \quad (10)$$

여기서  $M = cT \sum_{p=0}^{n-1} i(p)$  이고, 후항의 합계 항은 식 (8)과 같은 형태이므로 미소지연 필터로 구현할 수 있다. 식 (10)의  $M$  은 디지털 도파관에서 파동의 공간 위치를 나타내는 항목이므로 이는 지연의 위치에 의존적인 파라미터이며 또 한 파동의 속도에 의존적인 파라미터이다[10].

#### IV. 피치가 변하는 음 합성 결과

공간 기준 디지털 도파관 모델을 이용한 음합성 결과[6]와 본 논문에서 제안한 모델의 합성 결과를 비교할 것이다. 대상 음은 글리산도와 비브라토 음이다.

##### 4.1 글리산도

글리산도, 또는 슬라이딩이라고 하는 연주법에 의한 합성 음은 시작 음의 피치와 끝 음의 피치가 다르다. 음 피치를 올릴 수도 있고 내릴 수도 있다. 그림 4는 [4]에서 사용한 글리산도 파라미터 이다.

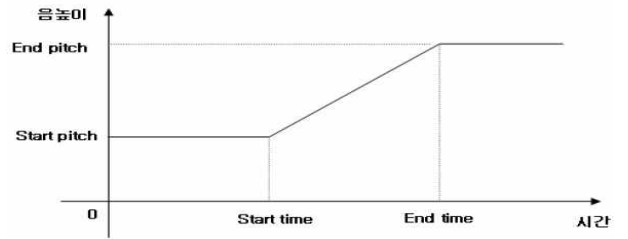


그림 3. 글리산도 (슬라이딩) 파라미터  
Fig. 3. Glissando (sliding) parameters

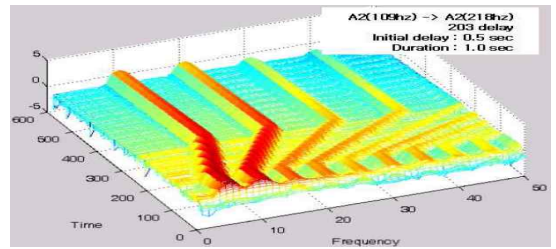


그림 4. 공간 기준 디지털 도파관 모델을 이용한 글리산도 (슬라이딩) 음 합성 결과  
Fig. 4. Synthesized glissando (sliding) sound by space-based digital waveguide model

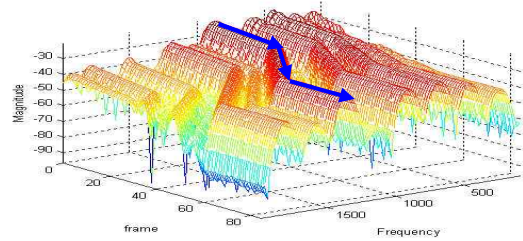


그림 5. 제안한 모델을 이용한 글리산도 (슬라이딩) 음 합성 결과  
Fig. 5. Synthesized glissando (sliding) sound by proposed model

그림 4는 공간 기준 디지털 도파관 모델을 이용한 글리산도 음 합성 결과이다. 지연 라인의 길이는 203으로 고정이고 파동의 속도를 2로 증가시켜 피치를 두 배로 올린 경우이다. 그림 5는 그림 4와 동일한 피치 변화를 제안한 모델로 합성한 음의 스펙트럼이다. 그림 4와 그림 5를 비교하여 볼 때 공간 기준 디지털 도파관에 의한 합성음의 경우 피치가 변하는 동안 배음의 크기가 현저히 줄어드는 반면 제안한 모델에서는 배음의 크기 변화가 잘 유지 되고 있음을 알 수 있다.

##### 4.2 비브라토

비브라토는 음의 높이가 반복되는 것으로 사인파, 톱니파, 구형파 등의 타입으로 표현할 수 있다. 그림 6은 사인파의 경우에 대한 파라미터이다.

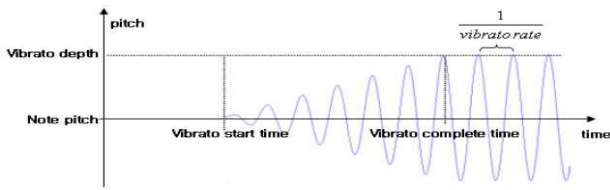


그림 6. 비브라토 파라미터  
Fig. 6. Vibrato parameters

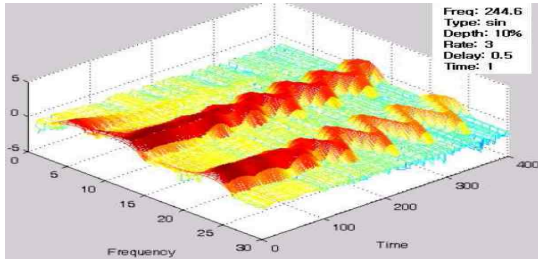


그림 7. 공간 기준 디지털 도파관 모델을 이용한 비브라토 음 합성 결과  
Fig. 7. Synthesized vibrato sound by space-based digital waveguide model

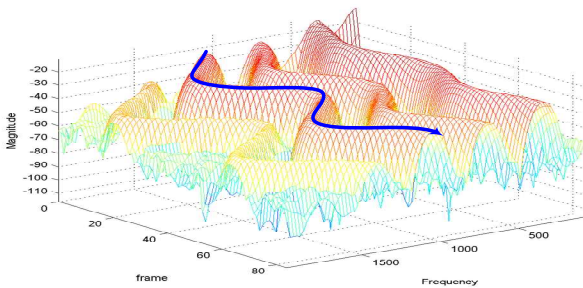


그림 8. 제안한 모델을 이용한 비브라토 음 합성 결과  
Fig. 8. Synthesized vibrato sound by proposed model

그림 7은 공간 기준 디지털 도파관 모델을 이용한 비브라토 음 합성 결과이다. 제안한 모델의 합성 결과인 그림 8은 그림 7과 매우 유사함을 알 수 있다.

## V. 결론

공간 기준 디지털 도파관 모델은 파동 방정식의 일반해를 이산 영역으로 옮기는 과정에서 공간 변수를 기준으로 정해 파동의 속도를 유동적으로 활용할 수 있도록 하였다. 파동의 상대 속도를 조절함으로써 파동이 단일 지연을 통과하는 시간을 달리하고 그 시간에 상응하는 피치를 변화시킬 수 있었다. 반면에 시간 기준 디지털 도파관 모델에서는 기본적으로 샘플링 시간이 단일 지연을 통과하는 시간으로 고정되어 있으므로 도파관 모델 자체만으로는 피치를 변화시킬 수 없다. 하지만 지연라인의 길이와 피치를 변화시킬 합성음에 대한 지연라인의 길이 비를 이용하여 도파

관 내에서 파동의 위치를 업 샘플링 또는 다운 샘플링을 통하여 피치가 변하는 음을 합성할 수 있었다. 주목할 만한 것은, 공간 기준 디지털 도파관 모델을 시간의 관점에서 바라보면 일정한 시간에 추출되는 샘플의 수가 변경되는 방식으로 샘플링 주파수를 시간에 따라 변경시키는 방식과 동일하다. [6]에서 공간 기준 디지털 도파관 모델에 의한 합성음은 페이지 보코더에 의한 합성음과 동일함을 보였다. 시간 기준 디지털 도파관 모델에서 사용한 피치 변환 방식은 [9]에서 제시하였으나 피치가 미세하게 바뀌는 모델에서 사용한 것으로 본 논문에서 사용한 것과 같이 피치를 크게 변화시키는 모델과는 차이가 있다.

본 논문에서 살펴본 두 가지의 디지털 도파관 모델은 각기 다른 관점에서 다른 방식으로 유도하였으나 지연 라인 내에서 단위 시간당 파동의 이동 거리를 변경한다는 점에서 동일한 개념이라 할 수 있다. 클리산도와 비브라토의 합성 결과를 바탕으로 볼 때 기타와 같이 강철 현을 가진 현악기나 관악기와 같이 많은 배음이 존재하는 악기의 경우 제안한 모델을 사용하는 것이 고주파 배음과 피치가 변하는 동안의 배음 크기를 유지할 수 있어 유리하고, 명주현이나 몸통이 작은 현악기의 경우에는 배음이 많지 않으므로 고주파가 빨리 감소하는 공간 기준 디지털 도파관 모델을 이용하는 것이 유리하다고 할 수 있다. 이와 같이 구현하고자 하는 악기에 따라 가장 적합한 도파관 모델을 선택하여 사용하면 될 것이다.

## 참고 문헌

- [1] *Information technology-coding of audiovisual objects-Part3: Audio*. ISO/IEC FCD Std. 14496-3, 1998.
- [2] K. Karplus and A. Strong, "Digital Synthesis of Plucked String and Drum Timbres." *Comput. Music J.*, Vol. 7, No. 2, pp. 43-55, 1983.
- [3] J. O. Smith, "Music Applications of Digital Waveguides," Technical Report STAN-M-39, CCRMA, Department of Music, Stanford University, Stanford, California, 1987.
- [4] Kyu-Nyun Kim, Myung-Soo Kang, "Physical modeling of plucked string instruments based on fixed spatial sampling interval," *Proceeding of KORUS 2000*, Vol. 2, pp. 1-6, 27 June - 1 July, 2000
- [5] Myung-Soo Kang, Kyu-Nyun Kim, "Delay characteristics and sound quality of digital waveguide model by spatial sampling," *Proceeding of KORUS 2003*, Vol. 2, pp. 92-96, 28 June - 6 July, 2003
- [6] 강명수, "공간 기준 디지털 도파관 모델을 이용한 클래식 기타의 클리산도 및 비브라토 음 합성", 박사학위논문, 울산대학교, 2003

- [7] J. O. Smith, "Physical modeling using digital waveguides," *Comput. Music J.*, Vol. 16, No. 4, pp. 74-91, 1992.
- [8] T. I. Laakso, V. Valimäki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay-Tools for fractional delay filter design," *IEEE Signal Processing Mag.*, Vol. 13, pp. 30-60, Jan. 1996.
- [9] T. Tolonen, V. Välimäki, and M. Karjalainen, "Modeling of tension modulation nonlinearity in plucked string," *IEEE Transactions on speech and audio processing*, Vol. 8, No. 3, May 2000
- [10] 조상진, 오훈, 정의필, "장력 변조 모델을 이용한 합성음의 피치 변환 구현", 한국음향학회 학술발표대회 논문집, Vol. 24, No. 1(s), pp.381-384, 2005년 5월



**조 상 진(Sangjin Cho)**

1997년 2월 울산대 전자공학과(공학사)  
 2002년 2월 울산대 전자공학과(공학석사)  
 2007년 2월 울산대 컴퓨터정보통신공학과  
 (공학박사)

2009년 3월 ~ 현재 울산대 전기전자정보시스템공학부  
 연구교수

※주관심분야 : 디지털신호처리, 악기음향, DSP 응용



**강 명 수(Myeongsu Kang)**

2008년 2월 울산대 컴퓨터정보통신공학부  
 (공학사)  
 2008년 3월 ~ 현재 울산대 컴퓨터정보  
 통신공학과 석사과정

※주관심분야 : 디지털신호처리, 악기음향, DSP 응용



**정 의 필(Uipil Chong)**

1978년 2월 울산대 전기공학과(공학사)  
 1980년 2월 고려대 전기공학과(공학석사)  
 1986년 1월 미국 Oregon State University  
 전기 및 컴퓨터공학과(공학석사)

1996년 12월 미국 Polytechnic University 전자공학과  
 (공학박사)

1986년 9월 ~ 1990년 12월 미국 Bruckner Mobile  
 Electronics, 전자 및 통신 엔지니어

1994년 1월 ~ 1996년 5월 미국 뉴욕 Polytechnic  
 University CATT 연구소 Research Fellow

1997년 3월 ~ 현재 울산대 컴퓨터정보통신공학부 교수

※주관심분야 : 디지털신호처리, 고장진단, 3D사운드