

비선형 왜곡법을 이용한 악기음의 합성

Musical Tone Synthesis using Nonlinear
Distortion Method

이 덕 수*, 성 평 모**

(Duck-Soo Lee, Keong-Mo Sung)

요 약

본 논문에서는 비선형 왜곡법을 사용하여 악기음을 합성하였다. 비선형 왜곡법은 기본적으로 하나의 정현파 발생기와 이에 의해 발생된 정현파에 인위적인 왜곡을 일으키는 증폭기로 구성된다. 출력파는 비선형 증폭기의 전달함수, 즉 파형 형성 함수를 변화시킴으로써 제어되는 배음을 갖는다.

파형 형성 함수는 녹음된 악기음을 분석하여 얻어진 각 부분음 크기를 바탕으로 체비셰프 다항식을 이용하여 얻어진다. 합성 시스템은 IBM-PC에서 수행되는 프로그램이며, 합성 결과는 파형과 스펙트럼 분석기에 의한 스펙트럼으로 보였다. 또한 합성 결과의 검증용 위해 전문가를 대상으로 시청 테스트를 수행하여, 합성음의 만족도를 평가하였다.

ABSTRACT

In this paper, musical tone, especially instrument tones are synthesized using nonlinear distortion technique. Nonlinear distortion is very simple but versatile method when you synthesize musical instrument tones. It basically consists of one sine oscillator and amplifier which makes distortion to input wave. Output wave has many harmonics that can be controlled by varying shaping function, which is the transfer function of nonlinear amplifier.

Shaping function is obtained from the analyzed harmonic amplitude data. Given harmonics amplitudes, Chebyshev polynomial is used to produce the shaping function that exactly makes the given harmonics at steady state. We constructed non-real time nonlinear distortion synthesizer program running at IBM-PC. To quantify the satisfaction of synthesized tones, listening test is carried out, and the result is presented.

I. 서 론

전자공학의 발달과 함께 악기음의 여러가지 특징을 실제 악기를 통하지 않고 전기를 이용하여 합성하는 연구가 활발히 이루어져 왔다. 즉 파형에 해당하는 전압을 인위적으로 발생시켜 스피커를 구동함으로써 음으로 들리게 하는 것이다. 전기를 이용하여

* 서울대학교 전자공학과 박사과정
** 서울대학교 전자공학과 교수

음파를 발생시킴으로써 악기의 음을 최초로 합성한 것은 1906년 커힐(Thaddeus Cahill)에 의해서였다. 그는 고압 전류로 발전된 정현파의 음들을 합하여 여러가지 음색을 만들어낼 수 있는 텔하모니움(Tel-harmonium)이란 전자 오르간을 제작하였다.

현재 많이 사용되고 있는 악기음 합성 방식으로는 악기 부분음 스펙트럼을 하나의 부분음에 하나의 정현파 발생기를 할당하여 이를 모아 음을 합성하는 가산 합성법(additive synthesis)과 풍부한 부분음을 포함한 일정한 파형을 발생시켜 이중 일부를 제거(filtering)하는 감산 합성법(subtractive synthesis)이 있으며, 최근에는 초우닝(J. M. Chowning)이 개발한 FM 합성법 등이 많은 음파 합성기에 채용되어 상용화되고 있다.⁽¹⁾

이러한 각 방식에 장단점이 각각 존재하나 본 논문에서는 원하는 스펙트럼을 자유자재로 발생시킬 수 있고 악기음의 자연스러움을 갖게하는 동적 스펙트럼(dynamic spectrum)이 가능한 비선형 왜곡법(nonlinear distortion synthesis)을 이용하여 악기음을 합성하였다. 비선형 왜곡법은 순수한 정현파를 왜곡을 일으키는 특성을 갖는 앰프에 가해줌으로써 많은 부분음을 갖는 음을 발생시키는 방법이다. 이 때의 왜곡의 정도, 즉 부분음의 크기는 체비셰프 다항식(Chebyshev polynomial)에 의해 완전히 제어되므로 원하는 부분음을 자유자재로 얻을 수 있다.⁽²⁾⁽³⁾

II. 비선형 왜곡법의 기본 원리

비선형 왜곡에 의한 음파 합성법(nonlinear distortion synthesis)은 음파를 전기적으로 합성함에 있어서 많은 수의 발전기를 사용하지 않고 풍부한 음색을 얻는 방법으로, 르 브룬(Le Brun)⁽⁴⁾과 다니엘 아르피브(Daniel Arfib)⁽⁵⁾ 두 사람이 독자적으로 연구하여 같은 시기에 발표함으로써 본격적으로 제시된 음파 합성 알고리즘이다. 자연의 음파와 같은 음을 합성하는 과정에서 가산 합성법(additive synthesis technique)보다 효율적인 음파 합성법을 찾기 위한 노력으로 개발된 많은 합성 기법중 FM 합성법 등과 함께 왜곡 기법(distortion technique)으로 불리우는 비선형 왜곡법은 합성음의 주파수 성분들이 동적으

로 변화하고(dynamic spectrum), 수행되는 연산의 양이 다른 방법, 특히 가산 합성법보다 훨씬 적게 되는 등의 많은 장점을 가진것으로 인정되고 있다.

비선형 왜곡법은 기본적으로 비선형적인 특성을 갖는 증폭기(amplifier)에 의한 음의 찌그러짐을 이용한 것이다. 가장 간단한 비선형 왜곡 음파 합성기의 블록도는 그림1과 같다. 비선형 왜곡법의 핵심이라 할 수 있는 파형을 왜곡시키는 블록을 파형 형성기(waveshaper)라 하고, 파형 형성기의 특성 함수를 파형 형성 함수(shaping function)라 한다. 문헌에 따라서는 파형 형성 함수가 전달함수(transfer function)라고도 칭해지지만, 일반적으로 전자공학에서 이 용어가 다른 뜻으로 사용되므로, 혼선을 막기 위하여 본 논문에서는 파형 형성 함수로 칭한다.

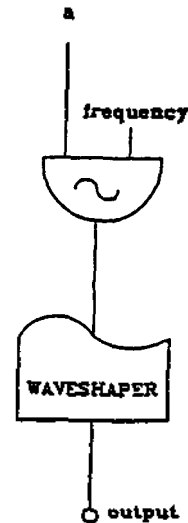


그림 1. 기본적인 비선형 왜곡 합성 시스템

파형 형성기(waveshaper)에의 입력은 순수한 정현파가 가해진다. 그림1의 정현파 발전기에 가해지는 입력은 정현파의 주파수 f 와 인덱스(index) a 가 있는데, 여기서 a 는 파형 형성기에의 입력 파형의 진폭에 해당하는 값으로서 동적 스펙트럼을 얻는데 필수적인 파라미터이다. 정현파 입력이 가해졌을 때 파형 형성기(waveshaper)를 통과한 출력 파형이 왜곡되는 현상을 그림2에 보였다. 그림 2(a)의 경우 선형적인

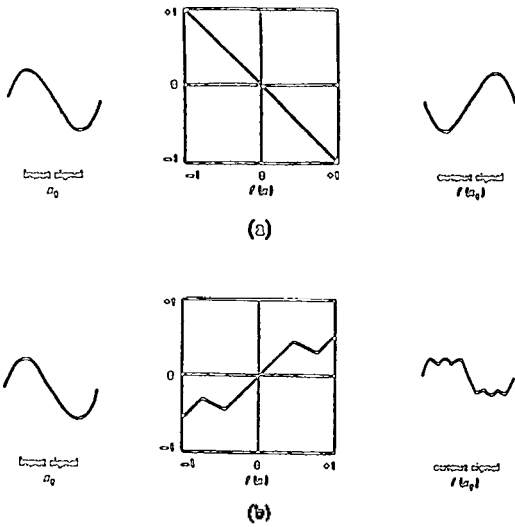
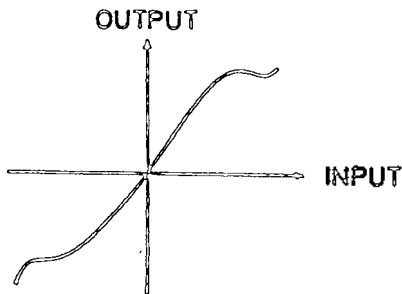
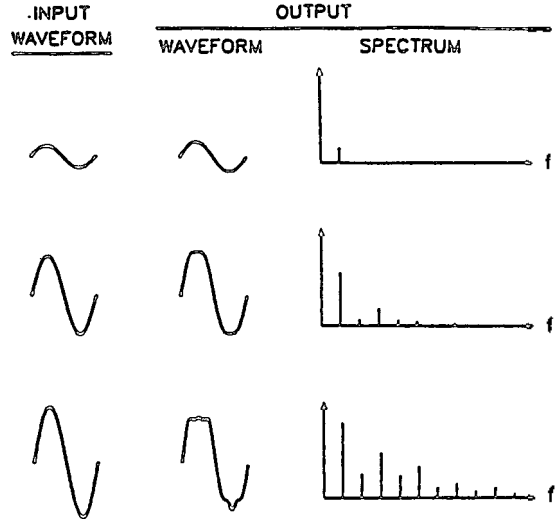


그림 2. 파형 형성 함수에 따른 입출력 관계
 (a) 사용된 shaping function
 (b) 입출력 파형과 출력 스펙트럼

파형형성 함수에 의한 출력 파형은 왜곡이 없이 진폭만 변화된 출력으로 나타난다. 그러나 그림 2 (b)의 경우, 출력파형은 비선형적인 파형 형성 함수에 의해 부분음 성분을 포함한 찌그러진 파형이 나타난다. 같은 파형 형성 함수를 사용하면서 입력 정현파의 진폭을 변화시키면 출력 파형은 입력 파형의 진폭에 따라 파형의 모양이 다양하게 변화한다. 그림 3(a)와 같은 파형 형성 함수를 가지는 파형 형성기(waveshaper)에 각각 다른 진폭의 입력을 가했을 때의 출력 파형과 스펙트럼을 그림 3(b)에 각각 보였다. 입력 정현파의 진폭이 작을 때는 파형 형성 함수의 선형 영역에서



(a) 사용된 shaping function



(b) 입출력 파형과 출력 스펙트럼...

그림 3. 입력 파형의 진폭의 따른 부분음 성분의 변화

진동하므로 부분음 성분이 없는 출력파가 얻어지지 만, 입력 정현파의 진폭이 커지면서 파형은 파형형성 함수의 비선형 영역을 통과하게 되어 파형이 찌그러 지면서 많은 부분음 성분을 갖게 됨을 알 수 있다. 그림 1의 기본 시스템에서 인덱스 a는 입력 파형의 진폭을 제어하는 파라미터로 사용되고 있으므로 이미 언급된 바와 같이, 인덱스 a를 변화시킴으로써 동적 스펙트럼을 얻을 수 있다.

Ⅲ. 합성시 고려되어야 할 요소

1. 부분음 스펙트럼

악기음의 음색과 물리적 파라미터 사이의 상관관 계를 규명하려는 노력은 오랫동안 많은 연구자들에 의해 이루어져 왔다. 많은 분석으로부터 얻어진 내용 들을 요약해 보면 악기음은 주기적인 파형을 가지 며, 음색은 파형과, 좀더 정확히 말해서 푸리에 (Fourier) 스펙트럼과 직접적인 연관을 갖는다는 것이다.⁶⁾

주기적인 파형이 스펙트럼상의 크기와 위상으로 분석이 가능하다는 것은 푸리에(Fourier)의 이론으로 이미 잘 알려져 있다. 최근에 초기 파천 및 스펙트럼

에 대한 연구가 이루어져 이의 중요성이 알려졌지만 과거의 연구는 정상상태의 음에 대한 연구가 대부분이었으며, 따라서 음색을 좌우하는 요소는 그 음의 스펙트럼상의 부분음의 크기인 것으로 보고되었고, ¹⁰⁾ 현재에도 이의 중요성은 타당한 것으로 인정된다. 음색을 규명함에 있어서 스펙트럼의 레벨에 대해서는 별 이견이 없으나 위상의 영향은 많은 논란이 되어 왔다. 일반적으로 음의 음향 법칙(Ohm's acoustical law)으로 알려진, 귀가 위상에 대해서는 귀머거리라는(The ear is phase deaf.) 주장에 의하면 두개의 같은 배음의 구성을 갖는 음이 서로 다른 위상 관계를 갖고 있다면 파형은 달라지지만 청취자는 두 음간의 차이를 느끼지 못한다는 것이다. 그러나 특별한 상황에서는 위상의 차이를 느낀다는 연구 결과도 알려져 있다.^{10b)} 결론적으로는 음색을 결정함에 있어서 위상이 영향을 미치기는 하지만 그 영향은 그다지 크지 않은 것으로 생각된다. 본 논문에서는 위상의 영향을 크게 다루지는 않기로 한다.

일반적으로 알려진 배음 성분과 음색과의 관계는

다음과 같다. 특징적인 배음 구성 중 첫째로 생각할 수 있는 것이 홀수 배음만 존재하는 악기의 경우이다. 이런 경우의 대표적 예가 클라리넷이다. 한쪽 끝은 막히고 다른 한 끝은 열린 관은 관의 공진 특성에 의해 홀수의 부분음만이 존재할 수 있게 된다. 클라리넷은 이러한 관의 경우와 흡사 하면서 마우스피스(mousepiece) 부분이 원통형 관의 형태에서 벗어나고 핑거 홀(finger hole)의 영향으로 어느 정도 다른 공진 특성을 보이지만 대략적으로는 홀수 부분음이 크게 가진된다고 말할 수 있다. 그림 4에 클라리넷 D4음의 스펙트럼을 보였다. 낮은 부분음에서는 홀수 부분음과 짝수 부분음의 크기 차이가 매우 크지만 높은 부분음에서는 이런 차이가 점점 없어진다. 이것은 앞서 말한 바와 같이 원통형 관의 공진에서 약간 벗어난 공진특성을 갖고 있기 때문이다.

두번째로 스펙트럼상의 포락선(envelope), 즉 포만트(formant)를 갖는 음의 경우를 생각할 수 있다. 현악기의 현을 손으로 뜯거나 활로 문질러 진동시키

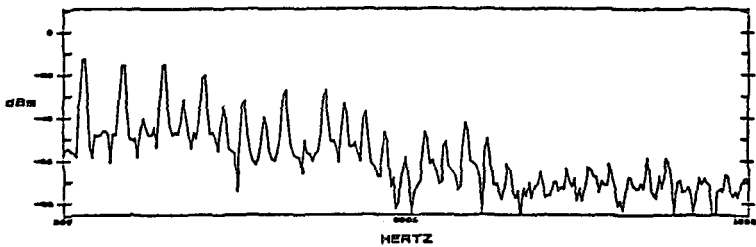


그림 4. 클라리넷 D4음의 정상상태 스펙트럼

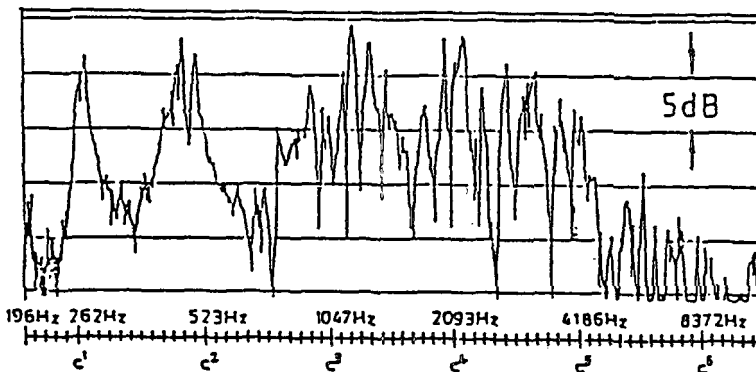


그림 5. 바이얼린 몸통의 공진 특성

면 이 진동은 브리지 (bridge)를 통하여 악기의 몸통에 전달된다. 전달된 진동에 의해 악기 몸통이 진동하면 이에 의해 악기 내부의 공기 및 외부의 공기가 진동하여 음파가 공기중으로 전달된다. 현악기의 몸통 및 브리지(bridge)는 복잡한 공진기 혹은 필터를 이루고 있으며 이 공진기는 상당히 많고 다양한 진동 모우드를 갖는다. 그림 5에 특정한 바이얼린의 브리지로부터 음향 출력시까지의 전달함수를 보았다. 진동 주파수가 전달함수의 높은 레벨 위치에 해당하면 그 부분음은 상대적으로 크게 소리나고 골짜기 부분에 해당하면 작은 소리가 된다. 바이얼린의 음색은 전달함수 혹은 공진 특성에 크게 좌우되며 훌륭한 음을 내는 바이얼린은 이러한 공진 특성이 귀에 듣기 좋도록 형성되어 있는 것이다. 따라서 현악기와 같이 특정한 공진 특성 혹은 포먼트(formant) 구조를 갖는 경우 이 점을 고려하여 음을 합성하여야만 성공적으로 음색을 재구성할 수 있다.

이밖에도 악기마다의 특징적인 스펙트럼이 존재하지만 일반적으로 관악기 등에서는 현존하는 악기의 정확한 스펙트럼을 흉내내는 것이 좋은 악기음을 합성하는데 필수불가결한 것은 아니다. 실험에 의하면 특정한 스펙트럼 데이터로부터 만들어진 파형 형성 함수 (shaping function)를 사용할 때 포락선(envelope)을 바꾸어 줌으로써 전혀 다른 악기의 소리를 만들어낼 수 있었다. 이의 원인은 파형 형성 함수(shaping function)는 정상상태의 부분음 크기를 파형 형성기(waveshaper)에의 입력이 최대일때 그대로 재현하여 주기는 하지만 입력의 크기가 최대가 아닐 때는 다른 모양의 스펙트럼을 만들기 때문이며, 또한 각 악기군마다 전형적인 진폭(amplitude)의 포락선(envelope)을 가져 이 포락선(envelope)이 악기를 구별하는데 매우 중요하기 때문이다. 이러한 점 때문에 각 악기의 스펙트럼이 그악기를 모사하는데 있어서 가장 좋은 모델이 되기는 하지만 실제로 합성 시스템을 구현할 경우 데이터 양을 줄이고 시스템을 간단히 하기 위해서는 전형적인 파형 형성 함수(shaping function)를 미리 저장하는 것도 좋을 것이다.

2. 초기 파형 및 스펙트럼

일반적으로 악기가 발음되는 과정에서 초기상태 즉 발음 시작부터 5-350ms 사이의, 초기 과도 상태(attack transient)라고 불러주는 상태는 매우 불안정하고 불규칙적이며 잡음성분 등을 포함하는 것이 보통이다. 어택(attack)부분은 짧은 시간에 일어나므로 과거의 연구자들은 분석하기가 매우 어려웠기 때문에 주로 정상상태에서의 스펙트럼이 악기음의 음색을 좌우하는 것으로 생각하였으나 최근의 연구 결과 초기 발음 특성이 악기음을 구별하는데 매우 중요한 파라미터임이 알려져있다.^[10] 인간이 음을 인식하는 과정에서 어택(attack)을 하나의 특성으로 받아들이는지, 아니면 귀나 두뇌가 이를 자세히 분석하는지는 확실히 알려져 있지 않지만 음을 구별하는데 있어서 매우 중요한 역할을 수행함은 잘 알려져 있다. 실험에 의하면 어택(attack)부분을 잘라내고 정상상태의 악기음만을 들려주었을 때 어떤 악기인지 구별할 수 없었다고 한다. 어택(attack)의 중요성은 특히관악기(목관, 금관악기)와 파이프 오르간등에서 두드러진다. 따라서 악기음 합성시 단순히 스펙트럼만을 그대로 재현하는 것만으로는 좋은 결과를 얻을 수 없으며 초기 특성을 반드시 고려해 주어야만 할 것이다. 비선형 왜곡법에서는 합성되는 악기음의 초기 특성을 그대로 모사하기는 어려우므로 동적 스펙트럼이 얻어지도록 하여 보다 실제 악기음에 유사한 음을 합성하였다.

3. 포락선(envelope)

각 음의 포락선(envelope)은 그 음의 특성에 부분음 스펙트럼만큼이나 큰 영향을 미친다.^{[11][12]} 실제로, 꼭 같은 부분음 스펙트럼을 갖는 음도 포락선(envelope)을 변화시킴에 따라 매우 다른 음색을 가진다.

악기음의 경우 종류에 따라 다양한 포락선(envelope)을 갖는 것을 살펴 볼 수 있지만, 악기음 합성사에 가장 널리 사용되는 것은 소위 ADSR 포락선(envelope)이라고 불러주는 것이다. 이는 그림 6과 같이 attack, decay, sustain, release로 구성되는 것으로 이의 조합에 의해 실존하는 악기음의 포락선(envelope)을 우수하게 모사할 수 있다.

포락선(envelope)이 음에 미치는 영향은 다음과

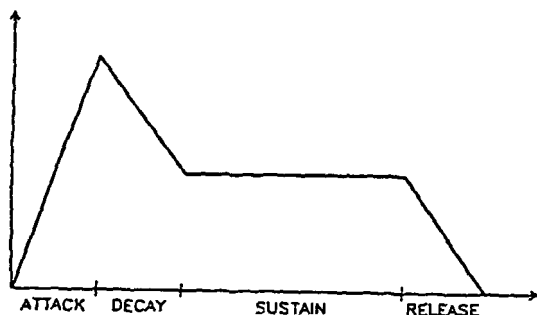


그림 6. ADSR 포락선

같이 요약될 수 있다.

1. 그림7(a)와 같은 포락선(envelope)은 타악기군의 소리를 모사한다. 이 때 음의 성분이 인하모닉(inharmonic)이면 북 등을 치는 소리가 되고, 하모닉(harmonic)이면 현악기의 피치카트(pizzicato, 손가락으로 현을 튕기는 주법)가 모사된다.
2. 그림7(b)는 길게 소리를 끄는 타악기군의 소리를 모사한다. 음의 성분이 인하모닉(inharmonic)이면 심벌(cymbal)류의 소리가 되고 하모닉(harmonic)이면 실로폰(xylophone)이나 비브라폰(Vibraphone) 등과 흡사해진다. 초기의 잡음(noise)에 의해 피아노와 같이 현의 타현에 의한 소리나 기타(guitar)현을 손가락으로 튕는 소리가 모사되기도 한다.
3. 그림 8(a)와 같은 포락선(envelope)은 목관악기군을 모사한다. 이 포락선(envelope)은 실제로 클라리넷음의 합성에 사용했던 것으로, 우수한 결과를 얻었다.
4. 그림8(b)는 금관악기군을 모사한다. 정상상태의 레벨보다 큰 레벨로 어택(attack) 부분이 가진 뒤 감쇠(decay)하는 이러한 모양은 금관악기에서 호흡을 붙여 넣는 순간에 해당한다.
5. 그림 9과 같이 느리게 변화하는 포락선(envelope)은 현악기군의 소리를 모사한다. 이 포락선(envelope)은 실제로 바이얼린 소리의 합성에 사용한 것으로 좋은 결과를 얻었다.

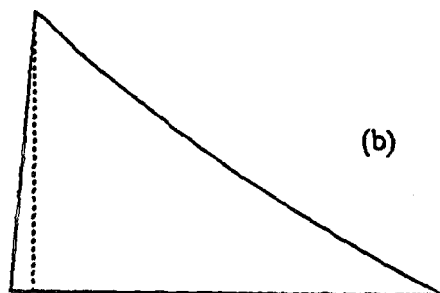
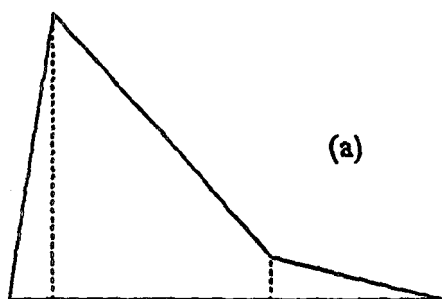


그림 7. 타악기군의 포락선
(a) 짧게 끝나는 타악기군
(b) 길게 울리는 타악기군

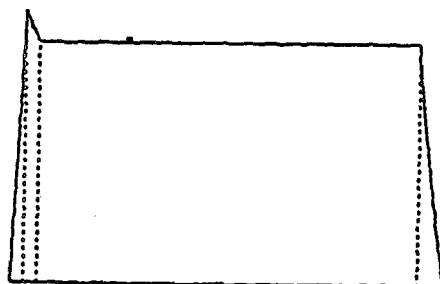
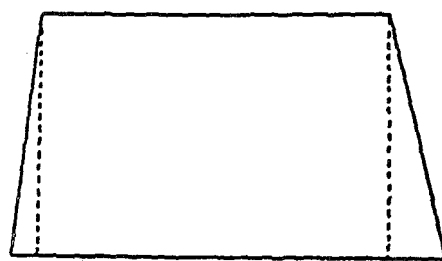


그림 8. 관악기군의 포락선
(a) 목관악기군
(b) 금관악기군

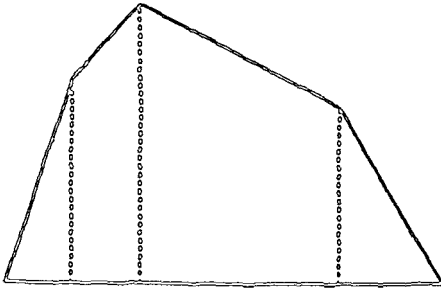


그림 9. 현악기군의 포락선

IV. 합성 과정

실험에 사용된 비선형 왜곡 합성 시스템은 다음 그림 10과 같은 구조를 갖는다. 비선형 왜곡법은 합성음의 스펙트럼 즉, 부분음의 크기를 제어하는 기법이므로 파형 형성함수(shaping function)를 정상 상태 스펙트럼 데이터로부터 계산한 후 이를 합성

시스템에 입력 데이터로 가해준다. 부분음 크기로부터 파형 형성 함수를 구하는 방법은 체비셰프 다항식 (Chebyshev polynomial)을 이용하여 각각 부분음을 급수의 계수로 변환한 후 이를 2048 바이트의 크기를 갖는 테이블로 만드는 것이다. 합성될 음의 음높이는 정현파 발진기의 입력 주파수로 결정되며 정현파 발진기의 출력을 포락선 발생기 1(envelope 1)의 출력에 곱해 줌으로써 동적 스펙트럼이 얻어지고 따라서 파형형성기의 출력 파형은 음색이 변화하게 된다. 파형 형성기의 출력은 다시 포락선 발생기 2(envelope 2)와 곱해져서 원하는 포락선을 갖게 된다.

그림 11에 실험에 사용된 장비들의 구성을 보았다. 전체 시스템은 크게 분석 파트와 합성 파트로 나뉘어지고, 분석 파트에서는 연주자들에 의해 실제로 연주된 각 악기음을 녹음하여 이를 주파수 분석기(spectrum analyzer)로 분석하는 과정을 수행한다. 분석을 통하여 얻어진 각 데이터는 합성 시스템

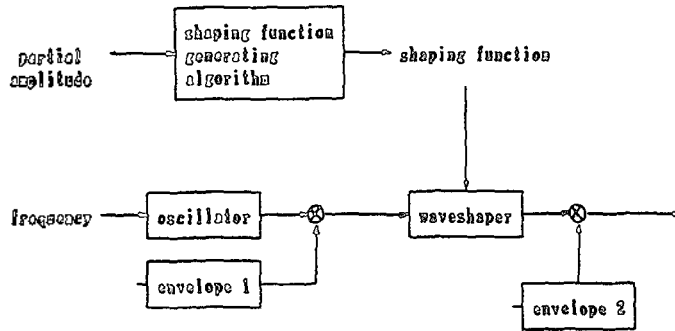


그림 10. 비선형 왜곡 합성 시스템

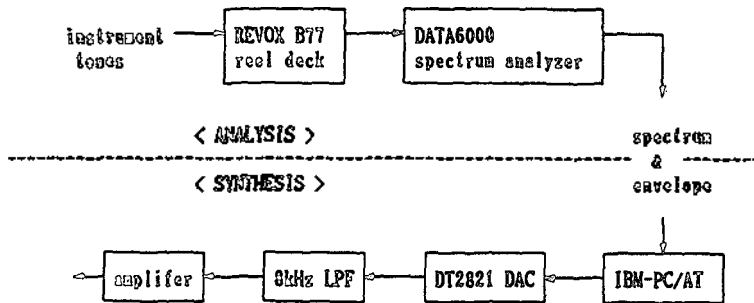


그림 11. 실험에 사용된 장비 구성

에 입력되는데 그림 10의 실제로 합성을 수행하는 블록은 IBM-PC/AT 상에서 수행되는 소프트웨어 시뮬레이터이며 C언어로 프로그램되었다. 합성을 데이타는 20kHz로 샘플링된 12 비트 바이너리(binary) 데이터 파일로 저장되며 D/A 변환기의 출력은 8kHz 대역폭을 갖는 저역 필터를 통과한다.

실제 악기음을 바탕으로 합성된 음을 만들어내는 과정을 그림 12와 같이 순서대로 나타내었다. 악기음은 각 악기별로 서울대학교 음악대학 해당 악기 전공생이 연주한 음을 녹음하였다. 연주는 해당 악기의 연주 음역내에서 연주 기법을 배제한 단음 연주를 기본으로 하여, 각 악기별로 사용되는 독특한 연구 기법을 사용한 음까지를 포함하여 녹음하였다. 단 본 논문에서는 비브라토(vibrato) 등의 연주 기법의 영향은 고려하지 않고 순수한 단음을 분석하

여 합성에 사용하였다. 각 음은 정상상태(steady state)를 갖도록 충분히 길게 연주하였으므로 분석시 얻어진 정상상태 스펙트럼을 사용하여 파형 형성 함수를 구하였다. 또한 시간축에서 본 파형의 크기 즉 포락선도 주파수 분석기를 통한 분석과정에서 얻어진다. 이러한 분석자료를 합성 시스템에 필요한 입력 데이터로 변환한 후 비선형 왜곡 합성 시스템에 입력하여 악기음을 합성한다.

합성된 음의 만족도는 원음과의 정량적 비교의 과정을 거치게 되는데 위상을 고려하지 않으므로 파형에 의한 비교는 판단의 기준을 제공하지 못하고 스펙트럼상의 비교에 주안점을 두었다. 또한 음색에 대한 판단은 주관적인 기준에 크게 좌우되므로, 시청 테스트를 통하여 합성음의 만족도를 평가하였다. 이러한 과정을 거친 후 평가 결과를 분석하여 이를 입력 데이터에 반영하는 과정을 반복하여 원하는

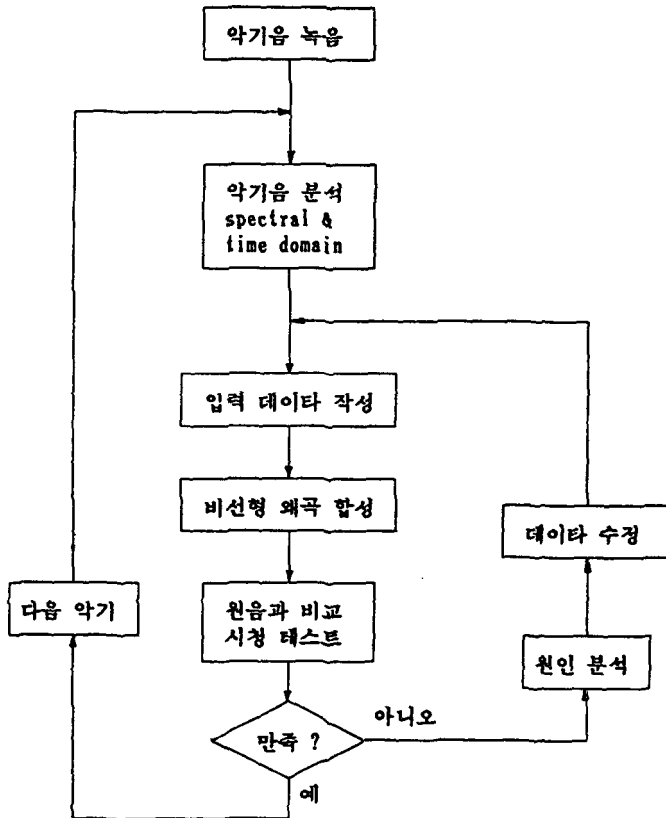


그림 12. 합성 과정 순서도

음색을 얻도록 하였다.

V. 합성 결과

1. 클라리넷 (Clarinet)

클라리넷은 플루트와 더불어 대표적인 목관악기로, 그 구조는 그림 13에 나타난 바와 같이 한 쪽 끝은 열려있고, 반대쪽 끝에는 마우스 피스(mouth piece)와 리이드(reed)가 달려 있다. 음을 발생할 때에는 마우스피스(mouth piece)쪽에 음압의 배(antinode)가 생기므로, 전체적으로 한 쪽 끝은 열리고 반대쪽은 닫힌 관과 비슷한 공진 특성을 가진다.

그림 14(a)에 실제 클라리넷 음의 초기 파형을, (b)에는 초기 스펙트럼, (c)와 (d)에는 각각 정상상태 파형 및 스펙트럼을 보였다. 초기 파형에서 어택(attack) 시간은 약 30ms 이고, 대략 10ms에 모든 배음이 다 발생되었음을 알 수 있다.

클라리넷 음의 합성은 실제 클라리넷 음의 분석

자료에 바탕을 두고 행하였다. 먼저, 그림 14(d)의 정상상태 스펙트럼으로부터 표1의 데이터를 구하였다. 그리고 이 표를 근거로 하여 만들어진 클라리넷 음의 합성을 위한 파형 형성 함수(shaping function)를 그림 15에 보였다. 또한 합성을 위한 포락선은 그림 16에 보였다. (a)는 파형형성기 앞 단에 걸리는 포락선으로서, 왜곡의 정도를 결정한다. 실제 악기의 파형 및 스펙트럼에서 보았던 바와 같이, 12ms만에 모든 부분음 성분들이 정상에 이르렀고, 마지막까지 음색의 변화는 없다. 그림 16의 (b)는 실제적인 출력의 진폭을 결정하는 포락선으로서, 어택(attack) 시간은 실제 악기보다 약간 긴 40ms를 사용하였다. 이는, 어택(attack) 부분에 직선적인 상향 곡선을 사용함으로써 지수 함수적인 곡선을 사용할 때보다 좀 더 급히 올라가는 경향을 고려하여 선택한 수치이다. 정상상태에서는 레벨의 변화없이 일반적인 목관악기의 포락선 모양을 유지하고 있고, 200ms에 걸친 release를 줌으로써 자연스런 음이 되도록 하였다.

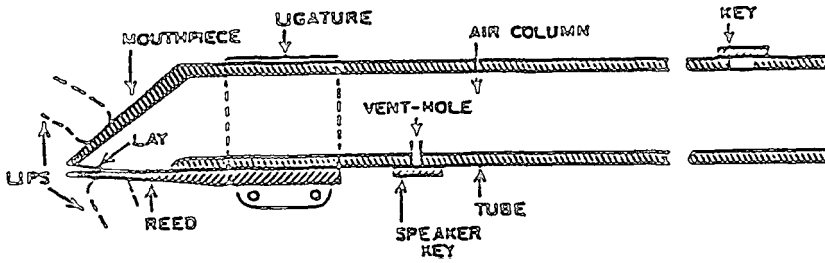
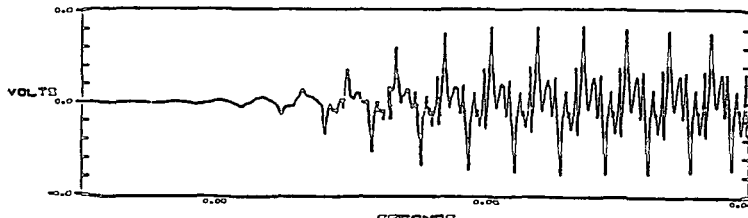
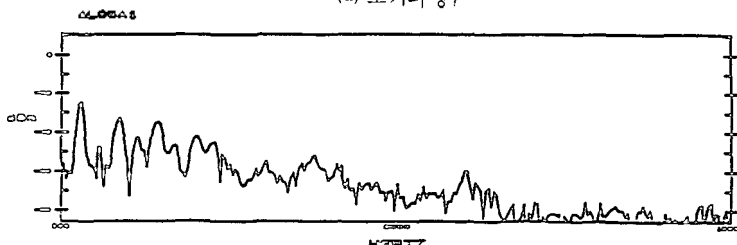


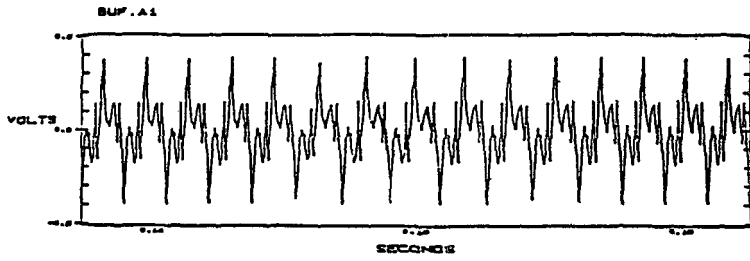
그림 13. 클라리넷의 구조



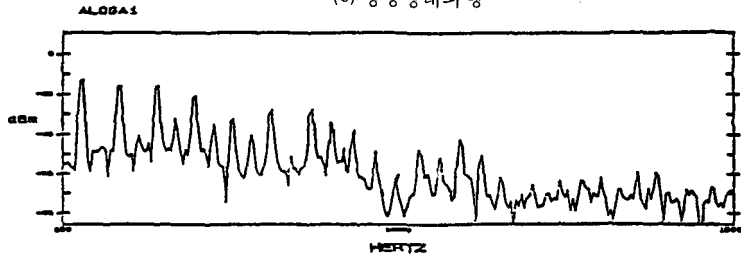
(a) 초기파형;



(b) 초기스펙트럼



(c) 정상상태파형



(d) 정상상태 스펙트럼

그림 14. 실제 클라리넷 음의 파형 및 스펙트럼

표 1. 클라리넷 D4음의 부분음 레벨

No.	레벨	No.	레벨	No.	레벨
1	63	11	59	21	32
2	34	12	19	22	19
3	65	13	48	23	21
4	34	14	44	24	13
5	70	15	36	25	19
6	55	16	32	26	23
7	59	17	21	27	23
8	46	18	32	28	15
9	62	19	32	29	21
10	53	20	25	30	11

*No: 부분음 번호, 레벨:[dB]

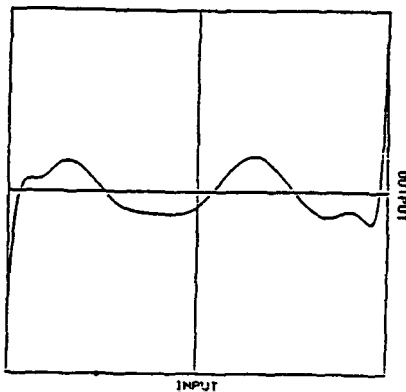
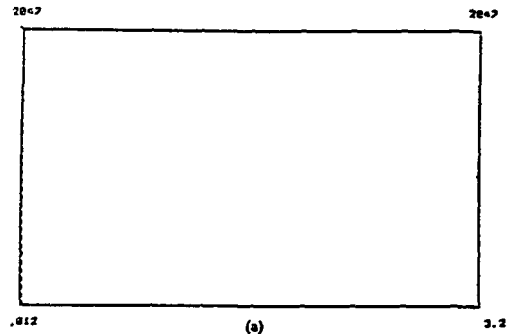


그림 15. 클라리넷의 파형 형성 함수



(a)



(b)

그림 16. 클라리넷의 포락선
(a) ENV1 (b) ENV 2

그림 17의 합성음을 그림 14의 실제 악기와 비교해 볼 때 많은 유사함을 볼 수 있으며, 합성된 음을 들어보았을 때에도 홀수 배음이 우세한 클라리넷의

특성을 느낄 수 있었다. 클라리넷 음은 홀수 배음이 뚜렷하게 우세한 특징을 가지고 있고, 잡음 등의 문제를 고려하지도 않아도 되기 때문에 상당히 합성

하기 쉬운 음이라고 할 수 있다. 따라서, 실제음의 분석에서 나온 자료를 별 수정없이 비선형 왜곡 시스템에 적용하여 좋은 결과를 얻었다.

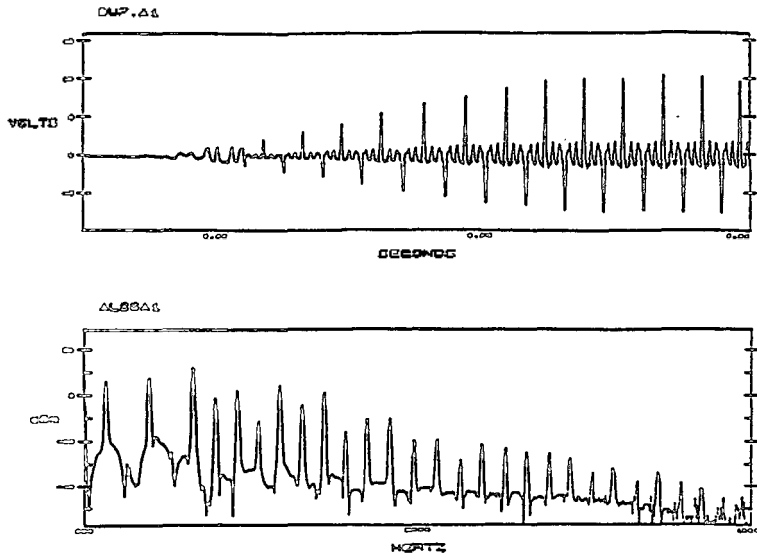


그림 17. 합성된 클라리넷 음의 파형 및 스펙트럼

2. 바이얼린 (Violin)

오케스트라 음악에서 가장 중요한 위치를 차지하는 악기인 바이얼린은 활이 현을 마찰시킴으로써 발생하는 뜸니파의 압력이 브리지(bridge)를 통해 전달되어 바이얼린 몸체를 진동시킴으로써 음을 발생시키게 된다. 바이얼린 몸통은 많은 수의 고유 공진 주파수를 가져 해당 주파수 성분으로 가진다면 큰 변위로 진동하고 따라서 음향 출력도 크게 되며 공진 주파수간의 골짜기부에 해당하는 주파수로 가진다면 음향 출력이 작게 된다. 가진된 뜸니파는 많은 부분음들로 구성되어 있고 이러한 부분음들은 각기 악기의 몸통이라는 복잡한 공진기를 거치므로 발생하는 음의 부분음들은 관악기의 등과 판이하게 다르게 된다. 즉, 인접부분음의 크기간에 상당한 레벨차를 보이며, 따라서 스펙트럼 포락선에 큰 변동(fluctuation)이 있게 되는 것이다. 이러한 특성은 현악기류의 독특한 요소로서 음색에 큰 영향을 미치

게 되며, 실험에 의하면 200-5kHz 대역에 20-30개의 공진 필터를 설치하여 뜸니파를 입력시킴으로써 바이얼린의 음색을 모사할 수 있었다.¹³⁾ 공진 특성의 산과 골짜기 비(ratio)는 약 10dB일 때 가장 좋은 음색을 얻었다.

바이얼린 A4음의 정상상태 스펙트럼 레벨을 표 2에 보였으며, 이를 바탕으로 만들어진 파형 형성 함수를 그림 18에 보였다. 바이얼린은 초기부터 많은 부분음 성분이 가진되어 음이 발생되므로 음색을 결정하는 ENV1은 그림 19(a)와 같은 간단한 형태를 사용하였다. ENV2는 그림 19(b)와 같은 데이터를 사용하였다.

합성된 음의 파형과 스펙트럼을 그림 20에 각각 보였다. 파형은 반주기를 기준으로 대칭을 이루고 있어 뜸니파와 무관해 보이니 스펙트럼은 현악기 특유의 스펙트럼 변동(fluctuation)을 보여 주고 있다.

표 2. 바이얼린 A4음의 부분음 레벨

No.	레벨	No.	레벨	No.	레벨
1	71	11	35	21	29
2	57	12	39	22	27
3	46	13	36	23	27
4	61	14	32	24	20
5	59	15	32	25	23
6	46	16	32	26	19
7	55	17	26	27	23
8	37	18	30	28	22
9	43	19	35	29	17
10	45	20	21	30	20

*No: 부분음 번호, 레벨: [dB]

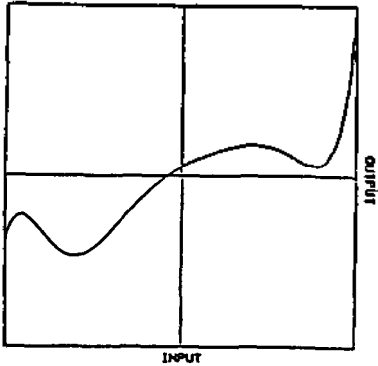


그림 18. 바이얼린의 파형형성 함수

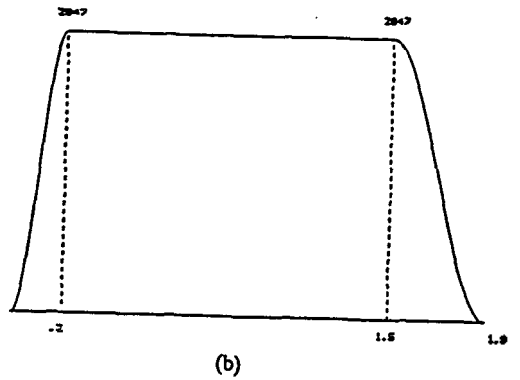
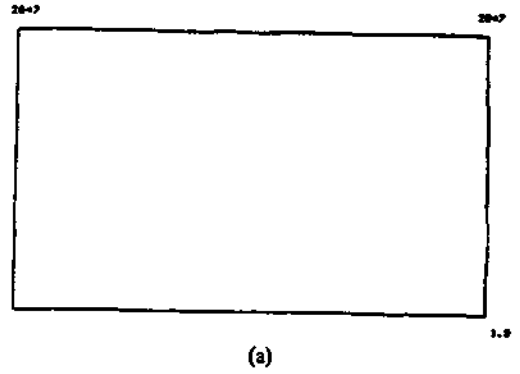
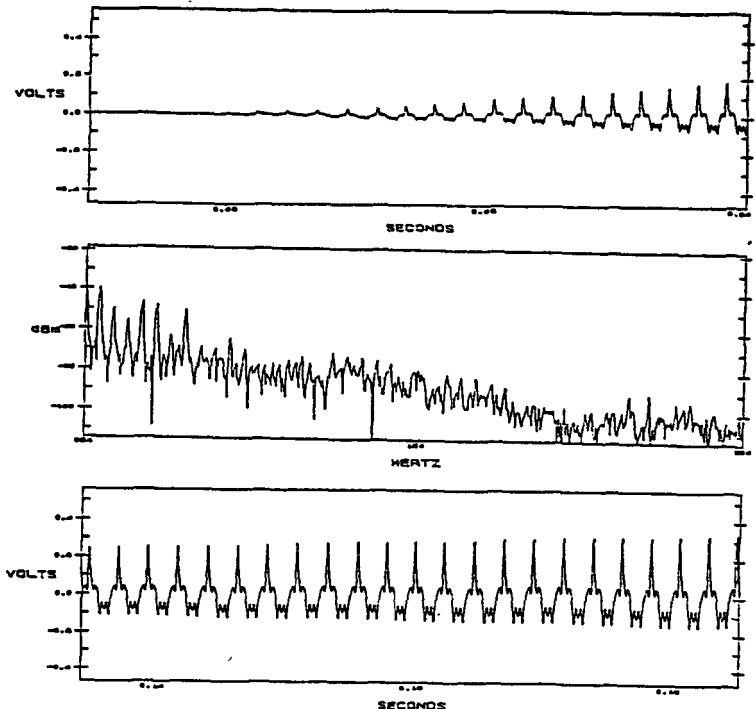


그림 19. 바이얼린의 포락선
(a) ENV 1 (b) ENV 2



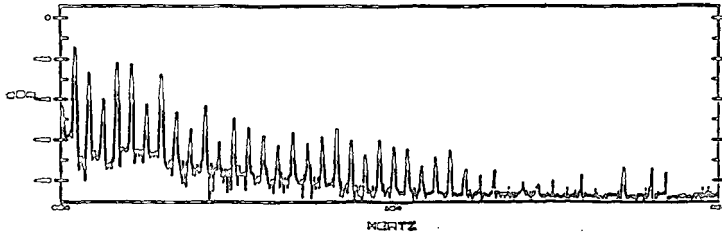
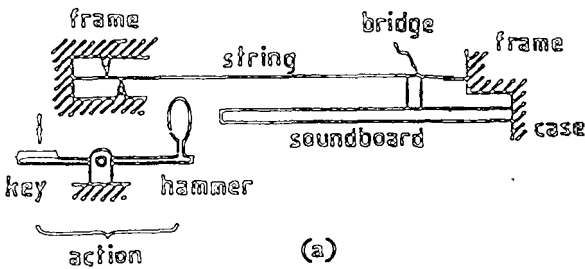


그림 20. 합성된 바이얼린 음의 파형 및 스펙트럼

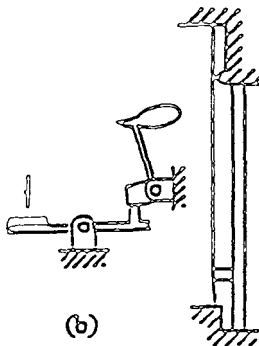
3. 피아노(Piano)

피아노는 그림 21과 같은 구조로 되어 있다. 프레임(frame) 및 브리지(bridge)에 고정되어 있는 현을 해머로 타격하여 현의 진동이 향판(sound board)에 전달되어 공기중으로 전파된다. 피아노의 현은 강성(stiffness)을 가지므로 양단을 고정시킨 후 진동시킬 때 고유진동수들의 기본 주파수의 배수를 이루지 않고 이보다 더 높은 주파수가 되어, 벗어나는

정도는 높은 주파수쪽으로 갈수록 더욱 심해진다. 따라서 피아노 음의 인하모니시티(inharmonicity)는 그림 22에 보인바와 같이 부분음의 번호가 높아짐에 따라 2차 함수적으로 증가하게 된다. 높은 음으로 갈수록 인하모니시티(inharmonicity)때문에 피아노의 원음과 같은 스펙트럼을 만들기 힘들어지는데 비선형 왜곡법은 원칙적으로 배음 성분만을 합성하는 음파 합성법이기 때문이다.



(a)



(b)

그림 21. 피아노의 간략화된 구조
(a) 그랜드 피아노 (b) 알라이트 피아노

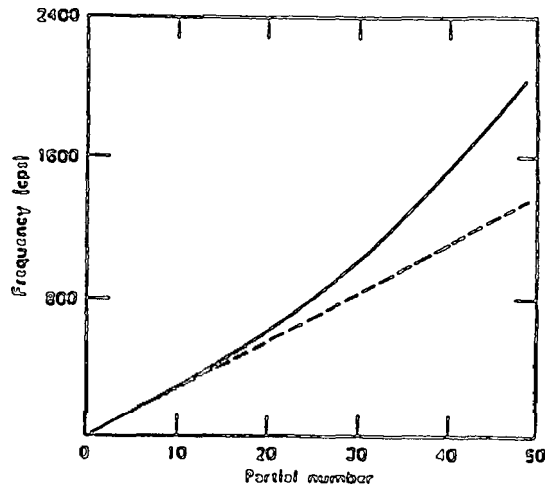


그림 22. 피아노의 인하모니시티

그림 23에 피아노 G1음의 파형과 스펙트럼을 보였다. G1 음과 같이 낮은 음을 우선적으로 합성 대상으로 선정한 이유는 위에서 설명된 바와 같이 피아노 음의 인하모니시티(inharmonicity)는 높은 음일수록, 높은 배음일수록 심해지므로 이의 영향을 최대한 배제시킬 수 있고, 피아노의 최저음부는 현이 하나

므로 여러개의 현에 의한 코러스(chorus) 효과를 고려하지 않아도 되기 때문이다. 그림 23의 스펙트럼으로부터 표 3과 같은 부분음 성분의 크기를 구하여 파형 형성 함수 계산의 입력 데이터로 사용하였다. 입력 데이터에 의해 구해진 파형 형성 함수를 그림 24에 보였다. 부분음의 개수가 많으므로 파형 형성 함수는 고차의 복잡한 함수를 이룬다. 피아노는 해머의 타격에 의한 빠른 어택 시간을 갖고 어택이 끝나면 대체적으로 지수함수적으로 감쇠하는 포락선을 사용한다. 그림 25에 합성에 사용된 ENV1과 ENV2

를 보였다. ENV1은 왜곡되는 정도와 동적 스펙트럼을 제어하는데 여기서는 특별한 의미가 없어 선형으로 감쇠하는 모양을 사용하였다. 실질적으로 중요한 ENV2는 지수함수적인 감쇠를 간단히 2개의 기울기가 다른 직선으로 근사하였다. G1등의 낮은 음은 60dB 감쇠 시간이 30초 정도로 매우 길지만, 실험의 편의를 위하여 7초만에 완전히 레벨이 0으로 떨어지도록 하였으므로 실제 피아노와는 차이가 있으나 합성 악기로는 문제가 없으며 음색을 구별하는 작업에는 큰 지장이 없었다.

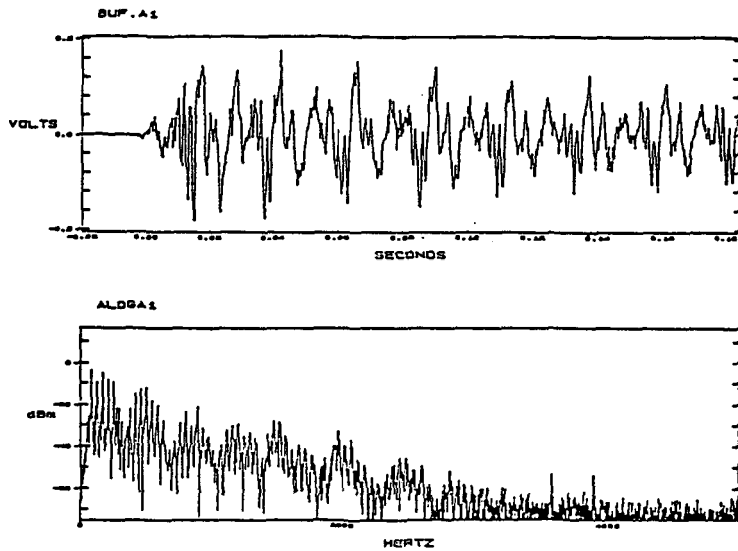


그림 23. 피아노 G1음의 파형 및 스펙트럼

표 3 피아노 G I 음의 부분음 레벨.

No.	레벨	No.	레벨	No.	레벨	No.	레벨	No.	레벨	No.	레벨
1	-100	11	51	21	44	31	38	41	28	51	12
2	55	12	53	22	47	32	32	42	19	52	14
3	52	13	48	23	35	33	24	43	11	53	21
4	55	14	47	24	31	34	26	44	22	54	21
5	53	15	41	25	22	35	40	45	32	55	29
6	52	16	19	26	28	36	40	46	29	56	21
7	41	17	30	27	37	37	35	47	20	57	15
8	24	18	39	28	41	38	24	48	24	58	17
9	43	19	30	29	40	39	29	49	22		
10	49	20	39	30	39	40	27	50	9		

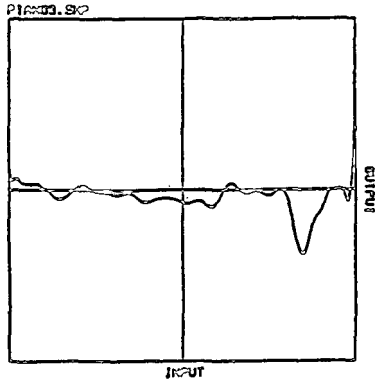


그림 24. 피아노의 파형 형성 함수

실제로 합성된 피아노음의 파형 및 스펙트럼을 그림 26에 보였다. 파형이 그림과 같은 모양을 갖는 것은 파형 형성 함수의 우측 끝에서 급격히 증가하고 그밖에는 대부분 평탄하게 진동하는 모양을 갖는 것에 기인한다. 파형으로 보아 원음과 별 유사성이 없는 것은 합성음 대부분이 갖는 성질이며, 위상에 따라 파형은 매우 다른 모양으로 보이지만 스펙트럼 성분의 전체적인 모양은 원음의 스펙트럼과 유사한 패턴임을 알 수 있다. 시청 결과, 원래의 피아노음과는 차이가 어느정도 존재하거나 피아노음으로 인식하는데는 문제가 없을 정도로 유사한 것으로 평가된다.

일반적으로 피아노는 샘플링 방법외에는 합성이 쉽지 않은 것으로 생각되고 있으므로 비록 한 음이

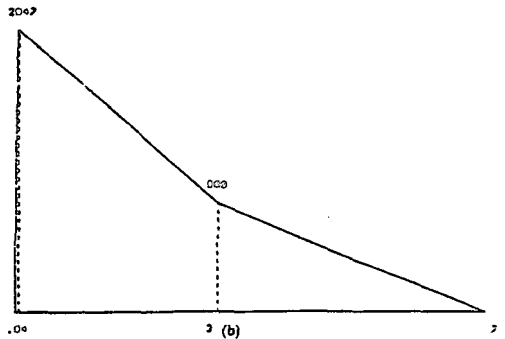
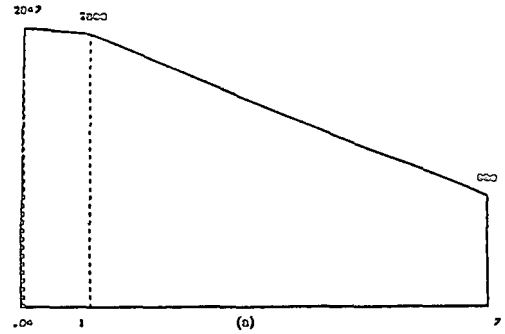


그림 25. 피아노의 포락선
(a) ENV1 (b) ENV2

지만 합성에 성공함으로써 매우 고무적인 출발점을 이루고 있다고 할 수 있다. 그러나, 높은 음의 합성은 비선형 왜곡 기본 시스템 만으로는 인하모니시티 (inharmonicity), 코러스(chorus) 효과 등의 분계점 때문에 어려우므로, 시스템 차원에서 연구가 더 이루어져야 하리라 예상된다.

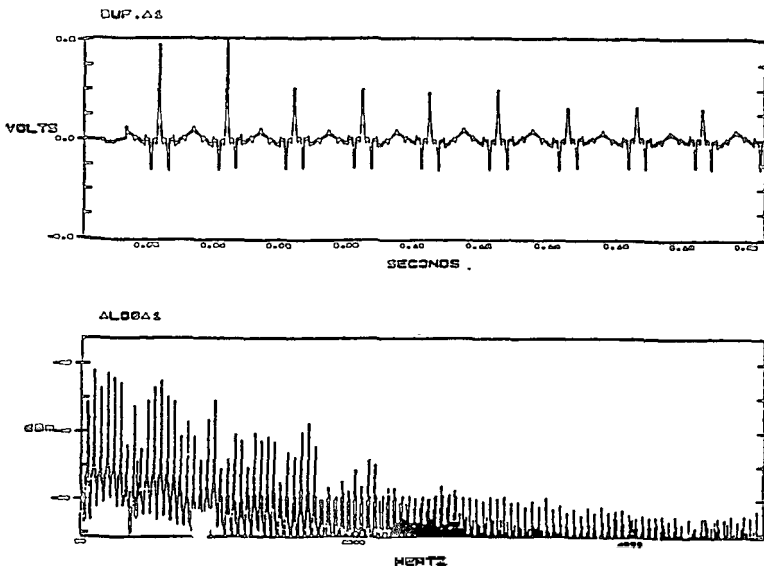


그림 26. 합성된 피아노 음의 파형 및 스펙트럼

4. 시청 테스트

지금까지 열거한 합성음을 포함해서 현재까지의 합성음을 정리하여 서울대학교 음악대학 학생들을 대상으로 시청 테스트를 하였다. 1학년에서 대학원 2학년까지의 작곡, 피아노, 성악, 현악 등 다양한 전공을 가진 학생을 대상으로 하였다.

시청 테스트는 2가지 방법으로 행하였다. 첫째는 녹음된 합성음을 들려주고, 4가지 악기의 보기를 주어 가장 비슷하다고 생각되는 한가지를 선택하게 하였고, 두번째 방법은 미리 목적인 악기의 이름을 알려 주고 난 뒤 합성음을 들려 주어 100점 만점의 평점을 쓰도록 하였다. 시청 테스트용 음을 녹음할 때에는 중간 중간에 실제 악기음도 넣어 효과를 측정하는데 도움이 되도록 하였다.

첫번째 방법의 테스트 결과를 요약하여 표4에 실었다. 전술한 바와 같이 4가지 보기 중에서 하나를

표 4 시청 테스트 첫번째 방법 결과

번호	악기명	인식률(%)	최고득점악기	비고
1	클라리넷	64	클라리넷	
2	피아노	60	피아노	
3	트럼펫	100	트럼펫	acoust.
4	플루우트	0	트럼펫	
5	클라리넷	57	클라리넷	vib.
6	오보에	100	오보에	vib.
7	트럼펫	13	프렌치호른	
8	바순	100	바순	acoust.
9	바이얼린	33	비올라	acoust.
10	첼로	93	첼로	
11	플루우트	33	트럼펫	vib.
12	바순	93	바순	
13	바이얼린	80	바이얼린	vib.
14	플루우트	60	플루우트	acoust.
15	트럼펫	50	트럼펫	vib.
16	오보에	87	오보에	
17	바이얼린	87	바이얼린	
18	더블베이스	100	더블베이스	
19	클라리넷	93	클라리넷	
20	오보에	80	오보에	vib.

*테스트 참가자수: 15명

고르게 하는 테스트이다. 표4에서 악기명은 그 음을 합성하면서 목적인 악기의 이름이고, 두번째란은 전체 응답자 수와 그 악기라고 응답한 사람 수의 비를 나타 내었다. 세번째 란의 인식률에는 그 비를 백분율로 표시 하였다. 최고 득점 악기명은 원래 합성시 목적했던 악기와는 다르다고 하더라도 가장 많은 응답자가 선택한 악기명을 기록하였다.

비고란에 Acoust, 라고 표시된 악기는 합성음이 아닌 실제 악기음을 들려준 경우이고, vib.라고 표시된 악기는 합성시 비브라트(vibrato)를 첨가한 합성음이다.

대체적으로 볼 때 클라리넷, 오보에, 바순 등 목관 악기와 바이얼린, 비올라, 첼로, 더블 베이스 등 현악기가 좋은 결과를 얻었다. 9번의 실제 바이얼린 음이 비올라로 인식된 것은 바이얼린 음역에서 너무 낮은 음을 선택해 들려주었기 때문이다. 표 5에 실제 악기음과 합성음의 인식률 평균과 실제 악기음의 평균 이상의 인식률을 나타낸 합성음의 수를 보았다.

표6에는 두번째 테스트 결과를 요약하였다. 악기명과 비고는 앞의 경우와 같고, 득점 평균란에는 100점 만점의 평점을 평균하여 기록하고 그 다음란에는 표준편차를 기록하였다. 이 테스트에 참가한 응답자는 16명이다.

표 5 평균 인식률

실제 악기의 평균 인식률	73.3%
합성된 악기의 평균 인식률	68.1%
실제 악기의 평균 이상을 득점한 합성 악기의 수	9

앞의 경우와 마찬가지로, 클라리넷, 오보에, 바순, 바이얼린, 더블 베이스 등이 70점 이상의 고득점을 얻었다. 반면에 플루우트, 트럼펫등은 50점 미만의 저조한 점수를 보였다. 실제 악기음의 득점 평균은 81점이었으며, 합성 악기음의 경우는 68점이었다. 이 경우 실제 악기음보다 높은 점수를 얻은 합성음은 모두 5개였는데 첫번째 테스트의 경우보다 저조한 것은 무슨 악기인지 알려주고 청취할 경우 청취

자가 더 결과를 비판적으로 보게 되기 때문이다.

표 6 시청테스트 두번째 방법 결과

번호	악기명	특정평균	표준편차	비고
1	클라리넷	57.5	10.8	
2	피아노	64.7	27.3	
3	트럼펫	91.9	4.3	acoust.
4	플루이트	45.0	17.6	
5	클라리넷	60.0	20.0	vib.
6	오보에	85.3	6.7	vib.
7	트럼펫	49.4	11.6	
8	바순	82.5	22.4	acoust.
9	바이올린	85.3	14.7	acoust.
10	첼로	67.2	23.0	
11	플루이트	45.0	15.7	vib.
12	바순	81.4	10.5	
13	바이올린	77.8	7.7	vib.
14	플루이트	63.8	15.3	acoust.
15	트럼펫	49.7	16.2	vib.
16	오보에	78.1	18.1	
17	바이올린	77.4	13.6	
18	더블베이스	80.5	11.1	
19	클라리넷	80.6	6.8	
20	오보에	83.1	10.3	vib.

V. 결 론

비선형 왜곡법은 가산 합성법 등에 비해 구조가 간단하면서도 손쉽게 동적 스펙트럼(dynamic spectrum)을 얻을 수 있기 때문에 근래에 들어서 주목받기 시작하는 파형합성 기법이다. 비선형 왜곡법은 악기의 발음과정과 유사한 과정을 거쳐 음을 발생시키므로, 자연스러운 즉 실제 악기와 유사한 악기음을 합성할 수 있다.

본 논문은 국내에 아직 잘 알려져 있지 않은 비선형 왜곡법의 원리에 대해 고찰하고, 비선형 왜곡법에 의한 음과 합성기를 프로그램화하여 몇가지 악기음을 합성하였다. 또한 합성된 음의 만족도를 평가하기

위하여 서울대학교 음악대학 학생들을 대상으로 시청 테스트를 실시하여 좋은 결과를 얻었다. 악기음 합성의 결과를 전문가를 대상으로 청취 테스트를 하여 그 성과를 평가하는 것은 음의 평가가 다분히 주관적이기 때문에 달리 객관적으로 검토할 기준이 모호한 점을 감안하면 매우 유용한 방법이라 할 수 있다. 그러나 본 논문에서 시행된 시청테스트는 청취자의 수가 그리 많지 않다는 약점을 갖고 있다.

비선형 왜곡에 의한 악기음의 합성이 성공적으로 이루어짐은 확인되었으나 이를 실제로 상용 악기음 합성기에 적용하기에는 아직도 해결되지 않은 점들이 많이 존재한다. 그 문제점들은 다음과 같이 요약될 수 있다. 첫째, 현재와 같이 하나의 파형 형성 함수(shaping function)로 합성하는 경우, 음높이의 변화에 따라 부분음 스펙트럼은 같은 패턴을 유지하면서 퍼지거나(음높이가 높아질 경우) 모이게 되므로 음색이 너무 규칙적이 된다. 이는 녹음된 음의 재생속도를 변화시킬 경우 일어나는 현상과 같다. 이 점은 하나의 파형 형성 함수를 사용하지 말고 음높이에 따라 파형 형성 함수를 바꾸어 주거나 동적으로 변화시켜 주면서, 필터(filter)등을 사용하면 어느 정도 해결될 수 있을 것이다. 둘째, 인하모닉(inharmonic) 스펙트럼을 얻는 방법으로 일반적으로 알려진 링 변조(ring modulation)를 사용하여 인하모닉 스펙트럼을 얻을 수 있을 것이다. 셋째, 본 논문에서 사용된 방법을 포함한 여러 기법들이 공통적으로 안고 있는 문제는 제어의 어려움이다. 즉 얻을 수 있는 스펙트럼은 이론적으로 추적하기에는 너무 복잡한 경우가 대부분이기 때문에 다분히 실험적인, 혹은 경험적인 면에 의존한다는 점이다. 넷째, 현재까지의 연구는 비선형 왜곡법의 기본적인 면만을 중심으로 이루어져 있으나 이를 실제 사용하기 위해서는 여러가지 음향 처리 기술을 적용시킨 시스템이 구성되어야 하므로 계속하여 이에 대한 연구가 진행 되어야 한다.

1. J.M. Chowning, "The synthesis of complex audio spectra by means of frequency modulation," J. Audio Eng. Soc. 21(7), pp.526-534, 1973.
2. M.Le Brun, "Digital waveshaping synthesis," J. Audio Eng. Soc. 27(4), pp.250-266, 1979.
3. D.Arifib, "Digital synthesis of complex spectra by means of multiplication of nonlinear distorted sine wave," J.Audio Eng. Soc. 27(10), pp.757-768, 1979.
4. D.Charles and T.A.Jerse, Computer Music, Schirmer Books, N.Y., 1985.
5. C.Roads, "A tutorial on nonlinear distortion or waveshaping synthesis," Computer Music Journal 3(2), pp.29-34, 1979.
6. A.de Bruijn, "Timbre-classification of complex tones," Acoustica 40, pp.108-114, 1978.
7. R.Plomp, Aspects of tone sensation, Academic Press, N.Y., 1976.
8. R.C. Mathes and R. L. Miller, "Phase effects in monaural perception," J.Acoust. Soc. Am. 19, pp78 0-797, 1947.
9. R.Plomp and J. M. Steeneken, "Effect of phase on the timbre of complex tones," J.Acoust. Soc. Am. 46, pp.409-421, 1969.
10. H.F. Pollard and E. V. Jansson, "Analysis and assessment of musical starting transients," Acoustica 51, pp.249-262, 1982.
11. J.Backus, The Acoustical Foundations of Music, W.W.Norton & Company, N.Y., 1977.
12. 성평모, 서양악기 및 국악기의 음향 분석, 서울대학교 생산기술 연구소, 1987.
13. M.V. Mathews and J.Kohut, "Electronic simulation of violin resonances," J.Acoust. Soc. Am. 53, pp.1 620-1626, 1973.
14. J.Beauchamp, "Brass-tone synthesis by spectrum evolution matching with nonlinear functions," Computer Music Journal 3(2), pp.35-43, 1973.
15. J.C.Risset and M.V. Mathews, "Analysis of musical instrument tones," Physics Today 22(2), pp.23-30, 1969.
16. M.V. Mathews, J.E.Miller and E.E. Jr. David, "Pitch synchronous analysis of voiced sounds," J. Acoust. Soc. Am. 33, pp.179-186, 1961.
17. 신디사이저 입문, 세광 음악 출판사, 1985.

▲이 덕 수



1965년 3월 30일생
 1983년 서울 성남고등학교졸업
 1987년 서울대학교 공과대학
 전자공학과 졸업
 1989년 서울대학교 대학원전
 자공학과 졸업
 1989년~현재 서울대학교 대
 학원 전자공학과박사
 과정 재학중

▲성평모 7 권 1 호 참조