

## 컬러이미지-소리 변환 시스템에 관한 기초연구

# A Basic Study on the System of Converting Color Image into Sound

김성일<sup>+</sup> · 정진승

Sung-Ill Kim and Jin-Seung Jung

경남대학교 전자공학과

Department of Electronic Engineering, Kyungnam University

### 요 약

본 논문은 컬러이미지에서 소리를 연상하는 인간의 공감각적 기능을 모방하는 지능로봇의 개발을 위하여 색-음 상호변환에 기초한 응용 시스템의 구축을 목표로 하고 있다. 그 첫 번째 단계로서 컬러 이미지를 소리로 변환하는 방법론에 관한 기초 연구 및 이에 기반한 기본 시스템을 구현하고자 한다. 본 연구에서는 컬러 영상이 갖고 있는 색상과 소리의 파동, 즉 상호 주파수 특성에서의 유사성을 이용하여 가시영역에서 가청영역으로 변환하는 방법을 제시하고, 컬러모델 변환 및 변환된 모델에서의 히스토그램을 이용하여 컬러 영상을 소리의 청각적인 요소로 변환 가능하게 하는 색-음 변환 방법을 제시하고 있다. 또한 본 연구에서 제시된 색-음 변환 방법을 토대로 Microsoft Visual C++을 이용하여 코딩하고 실제 기본 응용 시스템을 구현해 보았다. 그 결과 색채와 소리의 상호 연관성 즉, 색채의 색상(Hue), 채도(Saturation) 및 명도(Intensity)를 음의 높이(Fundamental Frequency), 하모닉(Harmonics) 및 옥타브(Octave)에 각각 대응시키고, 변환된 음향 요소를 Csound로 합성하여 웨이브 파일로 변환됨을 확인하였다.

**키워드 :** 이미지-소리 변환, 소리-이미지 변환, RGB컬러 모델, HSI컬러 모델, 히스토그램

### Abstract

This paper aims for developing the intelligent robot emulating human synesthetic skills which associate a color image with sound, so that we are able to build an application system based on the principle of mutual conversion between color image and sound. As the first step, in this study, we have tried to realize a basic system using the color image to sound conversion. This study describes a new conversion method to convert color image into sound, based on the likelihood in the physical frequency information between light and sound. In addition, we present the method of converting color image into sound using color model conversion as well as histograms in the converted color model. In the basis of the method proposed in this study, we built a basic system using Microsoft Visual C++(ver. 6.0). The simulation results revealed that the hue, saturation and intensity elements of a input color image were converted into F0, harmonic and octave elements of a sound, respectively. The converted sound elements were synthesized to generate a sound source with WAV file format using Csound toolkit.

**Key Words :** Image-to-sound Conversion, Sound-to-image Conversion, RGB, HSI Color model, Histogram

## 1. 서 론

과장과 진동수는 물리, 수학적으로 역비례 관계에 있어 수학적으로 상호변환이 가능한 물리량이다. 또한 파동으로 정의될 수 있으며, 파동의 특성인 진폭과 파장(또는 주파수)을 가진다. 이러한 파동의 특성은 가시영역에 있는 색채 신호와 가청영역에 포함되는 소리를 연결하는 근본적인 실마

리를 제공한다.

소리는 물체의 떨림에 의해 생성되고 그 떨림이 공기의 압력을 변화시킴으로서 전달된다. 공기나 물 등 그것을 구성하는 분자들은 서로 등거리를 유지하려는 속성이 있다. 그래서 어느 한 부분의 분자들이 이웃 분자보다 더 가까이 밀집되어 있으면 등거리 유지를 위해 반대 방향으로 이동하고 그것이 다시 이웃 분자들을 밀집시키게 한다. 이처럼 소리는 물체의 떨림과 압력 파동으로 구성된다.

색채는 형태와 질감 등 다른 시각적 요소와 결합하여 얼마나 조화롭고 멋스러운가를 표현하는 수단이면서 동시에 목적이 되고 있다. 소리의 대표적인 음악 역시 다양한 악기와 창법을 동원하여 보다 조화로운 음으로 감동을 전하려 한다. 따라서 소리에 의한 색채 연상과 색채에 의한 소리 연상은 색채와 소리의 관련성 및 공감각을 위한 다양한 시도들에 대한 동기가 되었다[1,2].

접수일자 : 2010년 2월 24일

완료일자 : 2010년 4월 5일

+ :교신저자

감사의 글 : 이 논문은 2009년 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (2009-0072992)

이처럼 색채와 소리는 우리에게 어떠한 느낌이나 영감을 주는데, 이러한 색채와 소리의 상호 공통점을 찾아, 오늘날 소리를 영상으로 영상을 소리로 변환하기 위한 다양한 시도 [3-10]가 있었다. 이러한 음-색, 또는 색-음 변환 기술은 공연장에서의 조명기술, 음악 플레이어에서의 영상표현, 그밖에 미디어 아트 등 많은 곳에서 적용, 응용되고 있다. 색채와 음악을 접목하기 위한 다양한 시도에 대한 문화적, 역사적 배경을 살펴보면 색채와 소리(또는 음악)는 공통점이 있다는 사실을 발견하게 된다. 과거 시도되었던 두 감각의 상호연결은 감성에 의존한 가정에 그쳤으나, 최근 과학적 근거를 바탕으로 한 체계적인 이론으로 응용시스템을 구축한 예들[3-7]을 볼 수 있다. 음향의 진폭과 파장을 빛의 진폭 및 파장과 대응시켜 상호 변환하는 방법과, 색과 음의 정보를 입력받아 주파수로 변환하고 입력받은 가시주파수를 기준주파수에 근거하여 주파수 변환공식에 의해 가청주파수로 변환 출력하는 시스템[8-10]도 발견할 수 있다.

본 연구의 기대 효과로서 인간의 공감각적 기능을 모방하는 새로운 형태의 지능로봇 개발, 스마트폰 및 디지털카메라 등의 휴대형 디지털 단말기의 새로운 응용솔루션 제공, 기존 온/오프라인 광고 산업에 있어 새로운 광고전달매체로의 응용, 시각 및 청각장애인을 위한 컬러영상인식 및 청각인식 보조 장치의 개발 등 실로 다양한 분야의 산업영역에 응용 가능할 것으로 기대한다.

색-음 변환 기반 감성시스템에 관한 기초 연구로서, 본 논문의 주된 작업은 색상정보와 소리정보를 이용한 색채와 소리의 물리적 파동원리의 공통점을 기반으로 상호 연관성을 찾아 색채와 소리라는 두 감각을 변환이 가능하도록 연결하는 보다 직관적인 방법을 찾는 것이다. 본 논문에서는 색-음 변환방법으로서, HSI 컬러모델의 색상(Hue), 채도(Saturation) 및 명도(Intensity) 성분의 각각의 히스토그램 값을 기반으로 최대값을 구하고 음의 3요소인 음의 높이 즉, 기본주파수(Fundamental Frequency), 하모닉(Harmonics) 및 옥타브(Octave)로 각각 매핑하여 음원을 합성하는 방법을 제안한다. 이를 기반으로 최종적으로 컬러 입력 영상에서 소리로의 색-음 변환을 가능케 하는 기본 변환 시스템을 구현하고자 한다.

## 2. 색-음 변환을 위한 기초 이론

### 2.1 색-음 변환을 위한 컬러모델의 표현

컬러모델(또는 컬러공간)은 일반적으로 컬러들과 다른 컬러들과의 관계를 표현하는 방법이다. 서로 다른 영상처리 시스템은 여러 이유 때문에 서로 다른 컬러모델을 사용한다. 예를 들면, 컬러로 된 그림을 출판하는 기업은 CMY 컬러 모델을 사용하며, 컬러 CRT 모니터와 컴퓨터 그래픽스 시스템들은 RGB 컬러 모델을 사용한다. 또한 색상, 채도 및 명도를 각각 다루어야 하는 시스템들은 HSI 컬러모델을 사용한다. 그림 1은 본 연구에서 기본적으로 사용하는 RGB 컬러모델의 Red, Green, Blue의 세 가지 요소와, HSI 컬러모델의 Hue, Saturation, Intensity의 세 가지 요소 각각을 보여주고 있다[11,12].

RGB 컬러모델은 컴퓨터 그래픽스 시스템의 설계를 간단하게 하지만, 모든 어플리케이션에 이상적이지는 않다. 일반적으로 R, G, B 각각의 컬러 요소들은 상호 관계가 너무 크며, 빛의 밝기에 대한 정보가 분리되어 있지 않고 빛의 밝기에 따라 물체의 색이 영향을 받기 쉬운 단점이 존재한다.

이러한 점 때문에 몇몇 영상처리 알고리즘들은 수행이 어렵다. 이 때문에 히스토그램 평활화와 같은 많은 처리 기술들은 영상의 명암도 요소만으로 진행된다. 그러므로 이러한 처리들은 색상, 명도 및 채도가 분리되어 표현되는 HSI 컬러 모델을 사용하는 것이 오히려 구현하기가 더 쉽다.

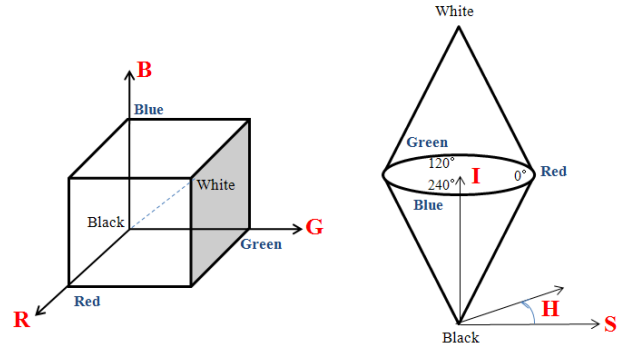


그림 1. RGB 컬러모델과 HSI 컬러모델

Fig. 1. RGB Color Model and HSI Color Model

HSI 컬러모델은 H는 색상, S는 채도, I는 명도를 각각 나타내며, 이 모델을 사용하면 어떤 구체적인 컬러를 만들기 위해 색을 조합할 필요가 없다. 바로 좌표축 H 자체가 색상을 나타내기 때문이다. 채도 S는 순수도를 나타내는 것으로서 원색에 어느 정도의 흰색이 혼합되었는지를 나타낸다. 그리고 I값을 조절하면 밝기를 변화시킬 수 있다. 이처럼 HSI 컬러모델은 색의 변환을 용이하게 해준다.

본 연구에서는 입력되는 RGB 컬러영상은 다음 식 (1)과 같은 변환식을 이용하여 HSI 컬러 모델로 변환한 다음 소리의 요소를 추출하는 과정을 거쳐 최종 음원을 생성하게 된다.

$$H = \cos^{-1} \left[ \frac{\frac{1}{2} [(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right]$$

$$S = 1 - \frac{3}{(R+G+B)} [\min(R, G, B)]$$

$$I = \frac{(R+G+B)}{3}$$

$$(If B > G, H = 360^\circ - H)$$

### 2.2 색-음 변환을 위한 소리의 표현

소리에는 그 소리의 특성을 구분할 수 있는 속성[13-15]이 있는데 음의 세기와 높이 그리고 음색이다. 예를 들면 고주파 대역의 고음으로 이루어진 음악은 밝게, 저주파 대역의 저음으로 이루어진 음악은 어둡게 느낀다. 또한 약기마다 구조나 재질에 의한 특징적인 배음 요소 때문에 다른 느낌을 준다.

음의 세 가지 요소 중 첫 번째로 기본주파수를 들 수 있다. 기본주파수는 음파의 기본음이 가지는 기본 주파수에 의해서 결정된다. 즉, 1초에 440번 진동하는 피아노의 A음은 440Hz의 음높이를, 1초에 880번 진동하는 음은 880Hz의 음높이를 가진 A음을 낸다. 따라서 주파수가 높은 음은 높

계, 주파수가 낮은 음은 낮게 들린다. 그림 2는 기본주파수로서의 12평균율(12음계 즉, 음악의 평균율 중에서 대표적인 것으로 현재 세계적인 표준 음률로 되어 있으며 옥타브를 12등분하여 1단위를 반음, 2단위를 온음으로 한 것)의 주파수 비율을 보여주고 있다. 여기서 보여 지는 바와 같이 12평균율의 각 음계는  $\sqrt[12]{2}$ 의 비율로 변화한다.

(12음계)

C	C#	D	E <sub>b</sub>	E	F	
1	$\sqrt[12]{2}$	$(\sqrt[12]{2})^2$	$(\sqrt[12]{2})^3$	$(\sqrt[12]{2})^4$	$(\sqrt[12]{2})^5$	
F#	G	A <sub>b</sub>	A	B <sub>b</sub>	B	C'
$(\sqrt[12]{2})^6$	$(\sqrt[12]{2})^7$	$(\sqrt[12]{2})^8$	$(\sqrt[12]{2})^9$	$(\sqrt[12]{2})^{10}$	$(\sqrt[12]{2})^{11}$	2

(주파수 비율)

그림 2. 기본주파수(F0)로서의 12음계의 주파수 비율  
Fig. 2. The frequency rate of 12 scales as a fundamental frequency(F0)

두 번째로, 음색에 영향을 끼치는 하모닉 성분은 음파를 구성하는 고조파의 구성 상태, 즉 하모닉 성분의 구조에 따라서 다르게 느껴진다. 즉, 외형상으로 매우 비슷한 악기라고 해도 고조파 배열과 크기가 다르면 두 악기는 서로 다른 음색을 가진다. 그림 3은 원음의 기본주파수와 하모닉 구조의 예를 보여주고 있다.

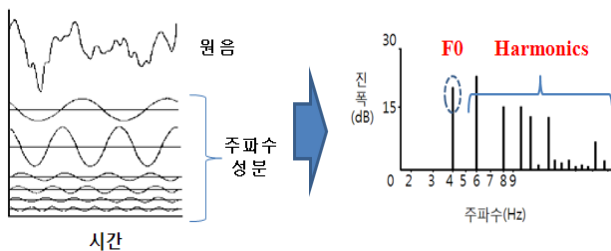


그림 3. 기본주파수(F0)와 하모닉 구조  
Fig. 3. The structure of F0(Fundamental frequency) and harmonics

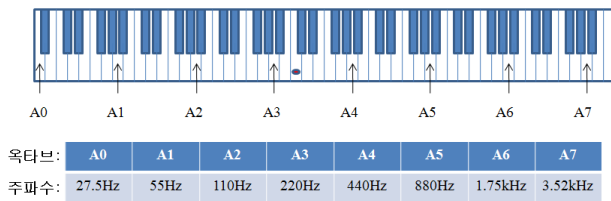


그림 4. 음계에서의 옥타브와 주파수 관계  
Fig. 4. The relation between octave and frequency in musical scales

마지막으로, 음 높이에서 발견되는 중요한 성질 중 하나가 옥타브다. 옥타브의 개념은 사람이 인지하는 독특한 현상으로 한 옥타브 차이의 음의 높낮이는 같은 음고로 물리적인 현상으로는 증명되지 않는 음악에만 존재하는 심리적인 현상 중 하나다. 진동수라는 성질로는 2도 관계나 3도

관계가 8도 관계보다는 수치상으로 가까운 값을 가지지만 사람은 8도의 음이 되면 다시 1도 음으로 되돌아오는 듯한 느낌을 받는다. 그림 4는 피아노 음계에서의 옥타브와 주파수의 관계를 보여주고 있다. 그림에서 알 수 있는 바와 같이 옥타브사이의 주파수는 배수의 관계를 지니고 있다.

### 2.3 음원 생성을 위한 Csound 개요

Csound[13,14]는 1985년 M.I.T. 교수인 Barry Vercoe에 의해 개발된 디지털 사운드 합성 프로그램으로서 원하는 모든 디지털 형태의 소리와 음악음향 효과들을 다양한 운영체제 속에서 모노(mono, 1채널) 혹은 스테레오(stereo, 2채널) 등 원하는 사운드 채널로 생성할 수 있는 텍스트 명령 기반의 소프트웨어 신디사이저 프로그램이다. Csound의 사용 환경은 MAC 혹은 PC에서 모두 사용 가능하며, 음악음향 및 전자음향 생성을 위한 데이터 처리 속도는 빠른 편이다.

이 프로그램은 컴퓨터 C언어를 기반으로 제작되었으며, ANSI-C언어를 기본으로 개발되어 각 컴퓨터 오퍼레이팅 시스템에서 컴파일만 하면 사용될 수 있다. 또한 Csound의 활용을 위한 C언어 선행학습은 불필요하며, Csound의 작동 원리를 이해하고 ORC(orchestra, 오케스트라) 파일과 SCO(score, 스코어) 파일 작성에 대한 충분한 이론학습과 경험이 뒷받침 된다면 훌륭한 음악음향을 제작할 수 있는 환경을 제공한다. 많은 기간을 거쳐 보안 수정된 Csound는 약 400개 이상의 신호처리 모듈을 이용하여 신디사이저를 만들 수도 있고, 멀티 이펙트 프로세서(multi effect processor)로도 이용할 수 있다. 이는 사용자가 원한다면 생성하고자 하는 모든 디지털 형태의 음악음향들을 제작할 수 있음을 의미한다.

그림 5는 Csound를 이용한 음원 생성과정을 보여주고 있다. 본 논문에서는 입력된 RGB 컬러 영상을 우선 HSI 컬러모델로 분리한 후, 색-음 변환 과정 즉, 컬러영상의 세 가지 특징 요소들을 소리의 요소들로 매핑시킨다. 최종적으로 ORC 파일과 SCO 파일을 자동 생성하고, Csound를 이용하여 음원을 합성, 출력한다.

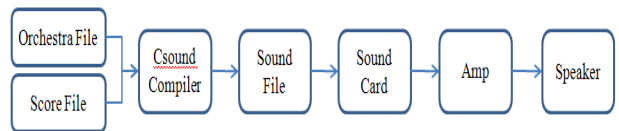


그림 5. Csound를 이용한 음원 생성과정  
Fig. 5. The process of generating a sound source using Csound

### 3. 색-음 변환의 원리

전자기파로 이뤄진 빛의 일부는 인간의 시각기관을 통하여 지각됨으로써 색채를 알 수 있다. 전자기파 중 380nm~780nm 파장대역, 즉 390THz(=3.9×10<sup>14</sup> Hz)~750THz(=7.5×10<sup>14</sup> Hz)의 빛이 어떤 에너지 분포를 갖느냐에 따라 다양한 컬러를 인지할 수 있다. 또한 가시광선 내에서는 파장에 따른 성질의 변화가 각각의 색깔로 나타나며 빨강색으로부터 보라색으로 갈수록 파장이 짧아진다.

공기의 진동을 매개로 전달되는 음향 또는 소리는 진동수(Hz)를 이용하여 표시하고, 사람이 들을 수 있는 청각 범

위, 약 20Hz~20kHz에 이르는 가청주파수 대역을 이룬다. 그림 6은 가청주파수와 가시주파수 대역사이의 주파수 범위를 각각 보여주고 있고, 본 논문에서는 가시주파수 영역의 컬러영상 이미지의 특징요소를 추출하여 가청주파수 대역의 특징요소로 각각 매핑한 후, 피아노(88 키보드)의 주파수 대역으로 제한하여 음원을 생성하였다.

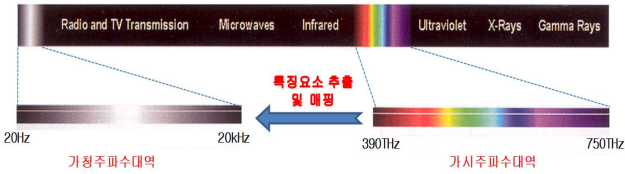


그림 6. 가청주파수와 가시주파수 대역사이의 주파수 매핑  
Fig. 6. The frequency Mapping between audible and visible frequency bands

색채와 소리는 물리적으로 공진, 증폭, 간섭, 상쇄 등의 파동의 속성을 공유하고 있고 도, 미 및 솔의 주파수 비율이 적색, 녹색 및 청색의 주파수 비율과 거의 일치한다. 그림 7은 이 원리를 이용하여 가시주파수 영역을 가청주파수로 주파수 변환하는 개념을 보여주고 있다. 즉, 가시영역의 주파수를 기본주파수로서의 12평균율로 변환하는 주파수 변환 공식을 유추할 수 있으며[8-10], 또한 가시주파수 대역(390THz ~ 750THz)은 거의 1 옥타브에 해당되고, 가청주파수 대역(20Hz ~ 20kHz)은 10 옥타브에 해당되므로 상호 대응관계를 찾을 수 있다[8-10].

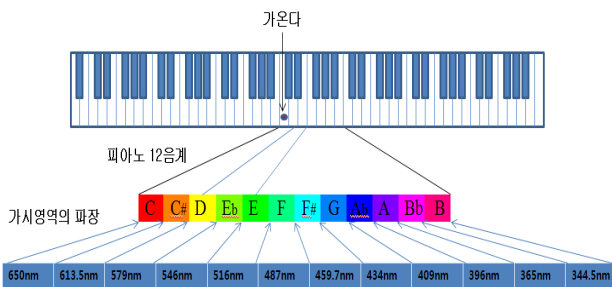


그림 7. 기본주파수(F0)로서의 12음계의 주파수  
Fig. 7. 12 kinds of frequencies in a scale as F0(fundamental frequency)

본 논문에서는 위에서 설명한 색-음 변환의 대표적인 기존 연구의 이론을 바탕으로, 그림 8에서 보여 지는 바와 같이 색채와 음의 대응관계를 정의하였다.

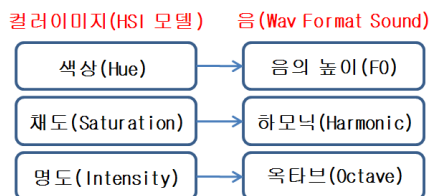


그림 8. 컬러이미지와 음의 대응관계  
Fig. 8. The mapping relation between color image and sound

색에는 그것을 구분하고 표현하여 전달할 수 있는 세 가지 속성이 있다. 색상, 명도, 채도가 바로 그것이며, 소리에도 그 음의 특성을 구분할 수 있는 속성이 있는데, 여기에는 음의 세기와 높이 그리고 음색 세 가지가 있다. 즉, 본 논문에서의 색-음 변환은 이미지의 색상, 명도, 및 채도를 소리의 피치, 옥타브 및 음색(하모닉 요소)에 각각 대응시키고, 대응 변환된 음의 요소를 합성하여 웨이브 포맷 파일로 변환한다. 그러므로, 본 논문에서는 색-음 변환 원리로서 주파수 변환공식을 유추하여 단순 변환시키는 것이 아니라, HSI 컬러모델에 기반한 컬러요소 각각에 음의 3요소를 매핑하는 방법에 주안점을 두고 있다. 또한, 음의 3요소를 추출하기 위하여 HSI 컬러 히스토그램을 이용하였다.

그림 9는 가시주파수 성분에서 HSI 컬러모델 중 색상 성분으로의 매핑, 그리고 음의 높이를 나타내는 음계 성분으로의 변환 과정을 보여주고 있다.

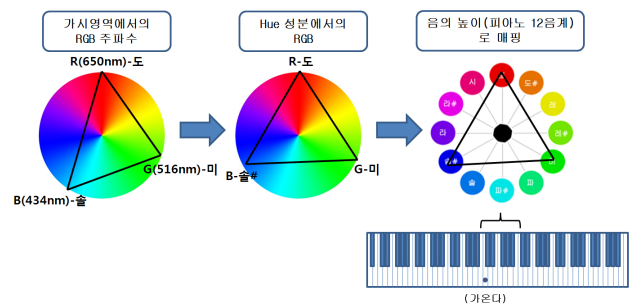


그림 9. 가시주파수에서 음계 성분으로의 변환 과정  
Fig. 9. The process of converting visible frequency into musical scale elements

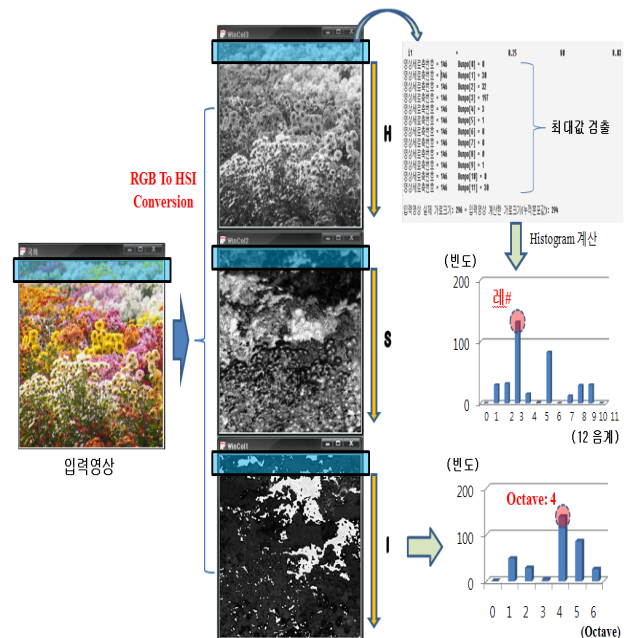


그림 10. HSI 컬러 히스토그램을 이용한 색-음 변환  
Fig. 10. Color image to sound conversion using HSI color histogram

가시영역의 대표적인 RGB 주파수는 도미솔의 주파수 비율과 일치하지만, HSI 컬러모델의 색상 성분은 RGB 상호



간에 120도의 관계를 가지고 있으므로 본 논문에서는 색상 성분을 30도씩 분리하여 피아노의 12음계로 매핑하였다. 즉, 기존연구[8-10]에서처럼 주파수 변환 공식에 따라 12음계의 주파수비와 일치하도록 색상과 음을 대응시켜 주지 않고, 색상을 동일 간격이 되도록 30도씩 그룹화하여 12음계로 대응시키는 비교적 단순하고 직관적인 방법을 이용하였다.

그림 10은 히스토그램을 이용한 색-음 변환 전 과정의 개념도를 간략히 보여주고 있다. 입의 컬러영상이 입력되면, 주어진 영상을 화소 단위로 분석하여 RGB 정보를 추출한다. 이렇게 추출한 RGB 정보는 HSI와 같이 색상, 채도, 명도를 기준으로 하는 좌표 체계의 정보로 바꾸어 준다. 변환된 HSI 값 중 색상 채널의 각각의 가로축 픽셀 값들을 기준으로 12개의 그룹으로 나누어 히스토그램을 구하고, 히스토그램의 최대값을 음의 높이로 대응시킨다. 채도 채널의 픽셀 값들은 동일한 방법으로 음색에 영향을 끼치는 하모닉 주파수 성분에 대응시켰다. 본 논문에서는 한 종류의 악기 사용을 전제로 하였으므로 6개의 하모닉 성분을 가진 하나의 음색으로 고정하여 이용하였다. 또한 명도 채널의 픽셀 값들도 동일한 방식으로 히스토그램을 구하고, 피아노(88키보드) 건반의 7개의 옥타브 주파수 성분에 대응시켰다.

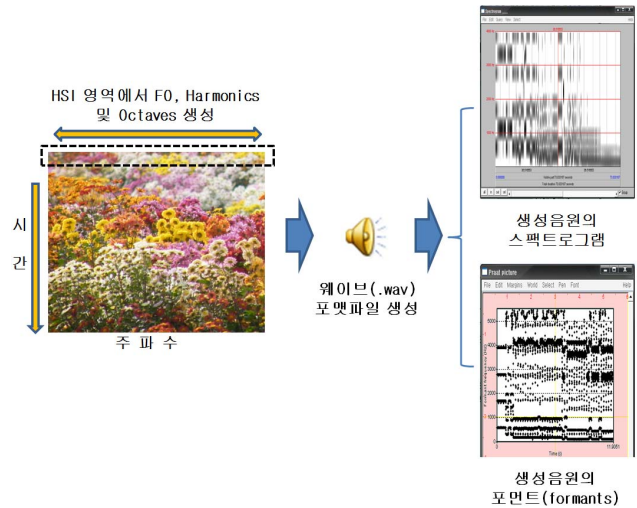


그림 11. 색-음 변환과정을 거쳐 생성된 음원 분석  
Fig. 11. The analysis of the generated a sound source through the procedure of converting color image into sound

#### 4. 실험 및 결과

본 논문에서 제안한 색-음 변환 원리를 실험, 평가하기 위해 실험 도구로는 Microsoft Visual C++, Csound를 사용하고 생성된 음원의 평가 도구로서 Praat를 사용하였다. 즉, RGB 컬러영상에서 HSI 컬러모델로의 변환과 색-음 변환 프로그램을 C++로 코딩하였고, 음원 생성을 위해 Csound를 사용하였다. 또한, 생성된 음원의 분석 툴로는 Praat를 이용하여 분석하였다. 그림 11은 히스토그램을 이용한 색-음 변환 과정을 거쳐 변환, 생성된 음향파일을 Praat를 이용하여 스펙트로그램(Spectrogram) 및 포먼트(Formants)를 추출한 결과를 보여주고 있다. 본 논문에서는 기본주파수 성분을 제외한 6 개의 하모닉 성분을 가진 하나의 음색으로 고정하였고, Praat를 사용한 스펙트로그램 및 포먼트 성분 추출에서 기본주파수 및 하모닉 성분을 확인할 수 있었다.

본 논문에서는 RGB 컬러모델에서 HSI 컬러모델로 변환한 후, H, S 및 I 각각의 채널의 세로 픽셀영역을 시간축으로 계산하였다. 그리고 색상 채널의 가로 픽셀영역은 히스토그램을 이용하여 최대값을 음의 높이에 해당하는 값으로 매핑하고, 명도 채널의 가로 픽셀영역에서 같은 방식으로 최대값을 구한 후 옥타브에 해당하는 값으로 매핑하여 기본적인 멜로디 성분을 합성하였다.

그림 12는 컬러이미지를 입력받아 생성된 음원으로부터 Praat를 이용하여 피치값을 분석한 결과를 보여주고 있다. 입력 컬러영상의 하늘 배경 부분(영역 1, 2)과 꽃줄기 부분(영역 4)은 비교적 비슷한 컬러영역이므로 거의 일정한 피치값을 유지함을 볼 수 있었다. 영역 3은 다양한 컬러영역이 혼합되어 있어 피치값도 그에 따라 변동됨을 알 수 있었다. 이와 같이 색상과 명도 채널로부터 최종 피치값을 추출하고, 이 값들은 입력되는 컬러 이미지의 색상의 변화에 따라 함께 변화함을 알 수 있었다.

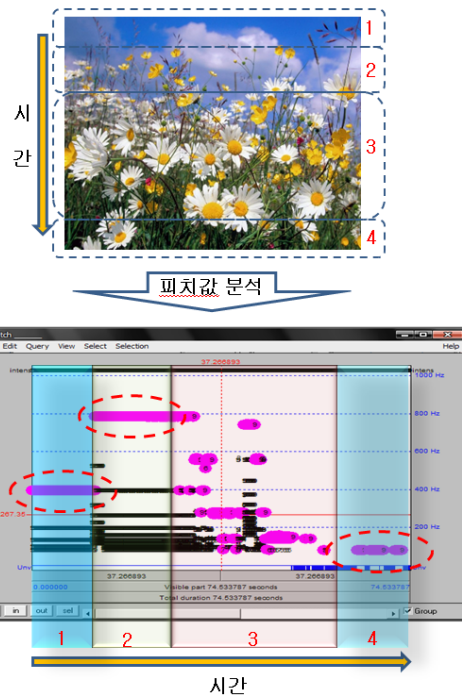


그림 12. 컬러 이미지로부터 생성된 음원의 피치값 분석  
Fig. 12. The analysis of the pitch of the sound source generated from an input color image

#### 5. 결 론

본 논문은 컬러이미지에서 소리를 연상하는 인간의 공간각적 기능을 모방하는 지능로봇의 개발을 궁극적인 목표로 하고 있고, 이를 위하여 색-음 상호변환에 기초한 응용 시스템을 구현하고자 한다. 그 첫 번째 기본연구로서, 컬러 이미지를 소리로 변환하는 방법론에 관한 기초 연구 및 이를

바탕으로 기본 시스템을 구현하였다. 실험 결과, 입력 RGB 컬러 영상을 HSI 영상으로 변환 분리함으로써 각각의 색상, 채도 및 명도 값을 음의 3요소인 음의 높이, 하모닉 및 옥타브에 각각 매핑시킴으로써 최종 음원이 생성됨을 확인하였다.

구현된 시스템의 문제점으로는 색-음 변환의 결과 음원의 소리는 뚜렷하게 출력 되지만 음악적인 요소가 배제되어 있어 음원만 듣기에는 부족한 점이 있었다. 앞으로의 연구 방향은 색-음 변환관계에 있어 컬러영상의 감성정보를 추출하여 음악적인 요소로 대응시키는 작업이 추가되어야 할 것이다. 이를 위해 영상의 3요소인 색상, 질감 및 형태의 정보추출이 필요하다. 또한 추출된 영상의 3요소를 음악의 3요소인 화음, 멜로디 및 하모니에 매핑시키는 최적의 방식을 찾는 것이 앞으로의 연구 방향이 될 것이며, 최종 시스템을 검증하기 위해 시청각 실험을 통해 입력 이미지와 출력의 음악적인 요소의 연관성 여부도 설문 조사할 예정이다.

### 참 고 문 헌

[1] 김민호, 정성환, 강민수, “음의 시각화와 그 표현의 경향”, *디자인學研究*, Vol. 16, No. 3, pp. 101-110, 2003.

[2] 박진희, “음악과 색채의 유기적 관계에 관한 연구: 모드를 중심으로”, *상명대 뉴미디어정보통신대학원*, 석사논문, 2005.

[3] N. Osmanovic, N. Hrustemovic, H. R. Myler, “A testbed for auralization of graphic art”, *IEEE Region 5, 2003 Annual Technical Conference*, pp.45-49, 2003.

[4] S. Matta, D. K. Kumar, X. Yu, M. Burry, “Discriminative analysis for image to sound mapping”, *Intelligent Sensing and Information Processing*, pp. 119-122, 2004.

[5] G. Bologna, B. Deville, T. Pun, “On the use of the auditory pathway to represent image scenes in real-time”, *Neurocomputing*, Vol. 72, No. 4/6, pp. 839-849, 2009.

[6] 이형민, *영상악기 만들기 1*, 마이크로소프트웨어, 10월호, pp. 234-239, 2006.

[7] 손우람, *영상악기 만들기 2*, 마이크로소프트웨어, 11월호, pp. 232-238, 2006.

[8] 김길호, 백정기, *사운드컬러하모니즘: 음악이 흐르는 컬러 배색 사전*, 임프레스, 2003.

[9] 김길호, *화성법을 이용하여 색과 음을 상호변환하고 색채를 조화하는 방법 및 장치*, 출원번호 10-99-34242, 1999.

[10] 김길호, *화성법을 이용한 음색변환표*, 출원번호 10-2001-0087651, 2001.

[11] 강동중, 하중은, *VISUAL C++을 이용한 디지털 영상처리*, 사이텍미디어, 2003.

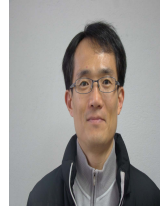
[12] 김용성, *VISUAL C++ 6 완벽가이드*, 영진닷컴, 2004.

[13] 이동재, *CSOUND*, 예당출판사, 2005.

[14] 석용환, *MIT CSOUND를 활용한 소리보기*, 천이미디어, 2006.

[15] 김유라, “소리의 표현 연구”, *서울여자대학교 대학원*, 석사학위논문, 1998.

### 저 자 소 개



#### 김성일(Sung-Il Kim)

1987.3~1994.2: 학사, 영남대학교 전자공학과

1994.3~1997.2: 석사, 영남대학교 전자공학과 대학원

1997.4~2000.3: 박사, (일본) Miyazaki University

2000.4~2001.3: 연구원, (일본)

NILS(National Institute for Longevity Sciences)

2001.4~2003.2: 연구원, (중국) CST(Center of Speech Technology), Tsinghua University

2003.3~2006.2: 전임강사, 경남대학교 전기전자공학부

2006.3~2010.2: 조교수, 경남대학교 전자공학과

2010.3-현재: 부교수, 경남대학교 전자공학과

관심분야 : 음성신호처리, 영상신호처리, 알고리즘

Phone : +82-55-249-2632

E-mail : kimstar@kyungnam.ac.kr



#### 정진승(Jin-Seung Jung)

1999.3~2007.8: 학사, 경남대학교 전자공학과

2008.3~2010.2: 석사, 경남대학교 전자공학과 대학원

관심분야 : 음성신호처리, 영상신호처리

Phone : +82-10-2912-5193

E-mail : jinseung80@naver.com