

# HSI 히스토그램에 기초한 이미지-사운드 변환

## Conversion of Image into Sound Based on HSI Histogram

김 성 일\*  
(Sung-Il Kim\*)

\*경남대학교 전자공학과

(접수일자: 2010년 10월 15일; 수정일자: 2010년 11월 22일; 채택일자: 2010년 12월 6일)

본 연구는 컬러이미지에서 특정 사운드를 연상하는 인간의 공감각적 기능을 모방하는 지능로봇의 개발을 최종 목표로 하고 있으며, 이는 컬러이미지와 사운드의 상호변환에 기초하여 이루어질 수 있다. 최종 목표의 첫 번째 단계로서, 본 연구는 컬러 이미지에서 사운드로의 변환을 이용한 기본 시스템의 구축에 초점을 맞추었다. 본 연구에서는 빛과 소리사이의 물리적 주파수 정보로부터 그 유사성에 기초하여 컬러이미지에서 사운드로 변환하는 방법을 제시한다. 컬러이미지에서 사운드로 변환하는 방법은 RGB-to-HSI 컬러모델 변환을 통한 HSI 히스토그램을 사용하고 Microsoft Visual C++을 이용하여 코딩함으로써 구현되었다. 두 가지 컬러이미지를 사용하여 시뮬레이션 실험을 하였고, 그 결과, 각각의 입력 컬러이미지의 색상 (Hue), 채도 (Saturation) 및 명도 (Intensity)를 사운드의 기본주파수 (F0: Fundamental Frequency), 하모닉 (Harmonics) 및 옥타브 (Octave)로 각각 변환한다. 제안된 시스템을 통하여 변환된 사운드 요소들을 Csound를 이용, 합성함으로써 웨이브 (wav) 파일 포맷 음원을 최종 생성하였다.

**핵심용어:** 이미지-사운드 변환, 사운드-이미지 변환, RGB 모델, HSI 모델, 히스토그램

**투고분야:** 뉴미디어 분야 (13)

The final aim of the present study is to develop the intelligent robot, emulating human synesthetic skills which make it possible to associate a color image with a specific sound. This can be done on the basis of the mutual conversion between color image and sound. As a first step of the final goal, this study focused on a basic system using a conversion of color image into sound. This study describes a proposed method to convert color image into sound, based on the likelihood in the physical frequency information between light and sound. The method of converting color image into sound was implemented by using HSI histograms through RGB-to-HSI color model conversion, which was done by Microsoft Visual C++ (ver. 6.0). Two different color images were used on the simulation experiments, and the results revealed that the hue, saturation and intensity elements of each input color image were converted into fundamental frequency, harmonic and octave elements of a sound, respectively. Through the proposed system, the converted sound elements were then synthesized to automatically generate a sound source with wav file format, using Csound.

**Keywords:** Image-to-sound Conversion, Sound-to-image Conversion, RGB model, HSI model, Histogram

**ASK subject classification:** New Media (13)

### I. 서론

파장과 진동수는 물리, 수학적으로 역비례 관계에 있고 수학적으로 상호변환이 가능한 물리량이다. 가시영역의 색채와 가청영역의 소리는 물리적으로 공진, 증폭, 간섭 및 상쇄 등의 파동의 속성을 공유하고 있으며, 도, 미, 솔의 주파수 비율이 적색, 녹색 및 청색의 주파수 비율과

거의 일치한다. 이처럼 색채와 소리는 파동으로 정의될 수 있으며, 파동의 특성인 진폭과 파장 (또는 주파수)을 가진다. 이러한 파동의 특성은 가시영역에 있는 색채 신호와 가청영역에 포함되는 소리를 연결하는 근본적인 실마리를 제공한다.

또한, 이들은 우리에게 어떠한 느낌이나 영감을 주는 데, 이러한 색채와 소리의 상호 공통점을 찾아, 오늘날 소리를 영상으로 영상을 소리로 변환하기 위한 다양한 시도 [1-9]가 있었다. 색채와 소리를 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, 과학적 근거를 바탕으로 한 체계적인 이론으로 응용시스템을 구축한 예들도 볼 수 있다. 흑백영상의 픽셀당 사운드 요소를 단순 맵핑한 시스템 [4]에서부터 음향의 진폭과 파장을 빛의 진폭 및 파장과 대응시켜 변환하는 방법 [5-6]과, 색과 음의 정보를 입력받아 주파수로 변환하고 입력받은 가시주파수를 기준주파수에 근거하여 주파수 변환공식에 의해 가청주파수로 변환 출력하는 시스템 [7-9]도 발견할 수 있다. 즉, 가시 영역의 주파수를 기본주파수로서의 12평균율로 변환하는 주파수 변환 공식을 유추할 수 있으며, 가시주파수 대역 (390 THz ~ 750 THz)은 거의 1 옥타브에 해당되고, 가청주파수 대역 (20 Hz ~ 20 kHz)은 10 옥타브에 해당되므로 상호 대응관계를 찾을 수 있다. 이처럼 소리에 의한 색채 연상과 색채에 의한 소리 연상은 색채와 소리의 관련성 및 공감각을 위한 다양한 시도들에 대한 동기가 되었다 [10-11].

전자기파 중 380 nm ~ 780 nm 파장대역, 즉 390 THz ( $=3.9 \times 10^{14}$  Hz)에서 750 THz ( $=7.5 \times 10^{14}$  Hz)의 대역을 가진 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 범위, 약 20 Hz ~ 20 kHz에 이르는 가청주파수 대역을 이룬다. 그림 1은 가청주파수와 가시주파수 대역사이의 주파수 범위를 각각 보여주고 있고, 본 연구는 인간의 시각기관을 통하여 지각되는 컬러이미지를 사운드로 변환하는 방법에 초점을 맞추고 있다.

본 연구는 특히 컬러이미지-사운드 변환에 기반한 감성시스템에 관한 기초 연구로서, 인간의 공감각적 기능을 모방하는 새로운 형태의 지능로봇의 개발을 최종목표로 하고 있다. 뿐만 아니라, 스마트폰 및 디지털카메라 등의 휴대형 디지털 단말기의 새로운 응용솔루션 제공,

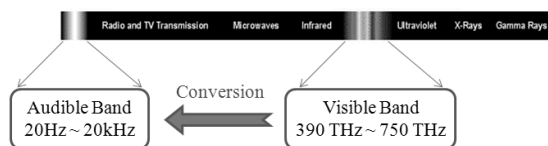


그림 1. 가청주파수와 가시주파수 대역사이의 주파수 변환  
 Fig. 1. The frequency Mapping between audible and visible frequency bands.

기존 온/오프라인 광고 산업에 있어 새로운 광고전달매체의 응용, 시각장애인을 위한 컬러영상인식 보조 장치의 개발 등 실로 다양한 분야의 산업영역에 응용 가능할 것으로 기대한다. 이러한 이미지-사운드 변환 기술은 실로 다양한 분야에서 적용, 응용될 수 있는 가능성을 가지고 있다고 말할 수 있으나 아직까지 확대된 연구의 진전뿐만 아니라 이 분야에 대한 인식조차 높다고는 할 수 없는 편이다.

본 연구의 주된 작업은 색상정보 [12]와 소리정보 [13]를 이용한 색채와 소리의 물리적 파동원리의 공통점을 기반으로 상호 연관성을 찾아 색채와 소리라는 두 감각을 변환이 가능하도록 연결하는 보다 직관적인 방법을 찾는 것이다. 위에서 언급한 기존 연구들은 영상과 소리사이의 한두 가지 요소들 사이의 단순매핑방법과 주파수 관계를 이용한 변환공식에 의한 매핑방법에 주안점을 두고 있다. 또한, 기존 지능형 로봇에 관한 연구 성과들은 카메라 및 마이크를 이용한 영상인식과 음성인식 등의 단순 결합에 그치고 있으며 이 둘의 결합을 통한 새로운 정보의 창출, 연상작용 등과 같은 복잡한 인간의 공감각적 기능을 가지기에는 아직 갈 길이 멀다고 할 수 있다.

본 연구에서는 컬러이미지-사운드 변환방법으로서, HSI 컬러모델의 색상 (Hue), 채도 (Saturation) 및 명도 (Intensity) 성분의 각각의 히스토그램 값을 기반으로 최대값을 구하고, 이들을 기본주파수 (Fundamental Frequency), 하모닉 (Harmonics) 및 옥타브 (Octave)로 각각 변환하여 음원을 합성하는 비교적 단순하지만 직관적인 방법을 제안한다. 이를 기반으로 최종적으로 입력 컬러이미지에서 사운드로의 변환을 가능케 하는 기본 변환 시스템을 구현하고자 한다.

## II. 컬러이미지-사운드 변환의 기초 이론

컬러들 사이의 관계를 표현하는 방법인 컬러모델 (또는 컬러공간)에는 현재 여러 가지 모델이 사용되고 있으나, 본 연구에서 사용된 대표적인 두 가지의 컬러모델인 RGB 및 HSI 컬러모델을 소개한다. 그림 2는 본 연구에서 기본적으로 사용하는 RGB 컬러모델의 Red, Green, Blue의 세 가지 요소와, HSI 컬러모델의 Hue, Saturation, Intensity의 세 가지 요소 각각을 보여주고 있다.

빛의 삼원색을 이용한 RGB 컬러모델의 경우, 컴퓨터 그래픽스 시스템의 설계를 간단하게 하지만, 모든 응용 어플리케이션에 이상적이지는 않다. 일반적으로 이 모델

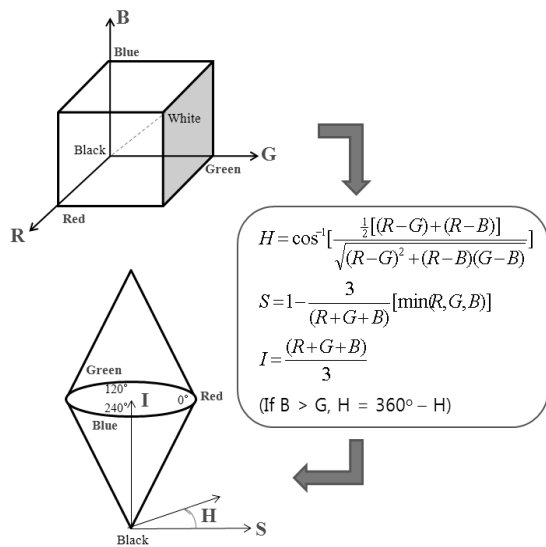


그림 2. RGB 컬러모델과 HSI 컬러모델  
Fig. 2. RGB Color Model and HSI Color Model.

의 단점으로는 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 점을 들 수 있다. 이 때문에 몇몇 영상처리 알고리즘들은 수행이 어렵고, 따라서 히스토그램 평활화와 같은 많은 처리 기술들은 영상의 명암도 요소만으로 진행된다. 그러므로 이러한 처리들은 색상, 명도 및 채도가 분리되어 표현되는 HSI 컬러 모형을 사용하는 것이 오히려 구현하기가 더 쉽다.

인간의 색 인지 방법에 기반을 둔 HSI 컬러모델의 경우, H는 색상, S는 채도, I는 명도를 각각 나타내며, 본 모형을 사용하면 어떤 구체적인 컬러를 만들기 위해 색을 조합할 필요가 없다. 좌표축 H 자체가 바로 색상을 나타내기 때문이다. 채도를 나타내는 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 RGB 컬러모델 기반 입력영상은 그림 2와 같은 변환식을 이용하여 HSI 컬러 모델로 변환한 다음, 색상, 채도 및 명도에 해당하는 소리의 요소, 즉 기본주파수, 하모닉 및 옥타브로의 대응 변환을 시도하였고, 최종 음원을 합성, 생성하였다.

사운드의 특성을 구분할 수 있는 속성에는 대표적으로 음의 높이, 음색, 그리고 음의 크기를 들 수 있다. 사운드의 첫 번째 요소로서 음의 높이, 즉 기본주파수를 들 수 있는데, 이는 음파의 기본음이 가지는 기본 주파수에 의해서 결정된다. 즉, 1초에 440번 진동하는 피아노의 A

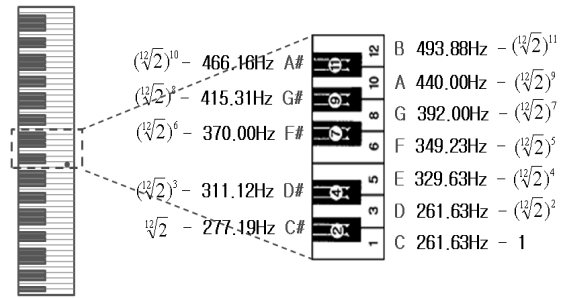


그림 3. 기본주파수로서의 12음계의 주파수 비율  
Fig. 3. The frequency rate of 12 scales as a fundamental frequency (F0).

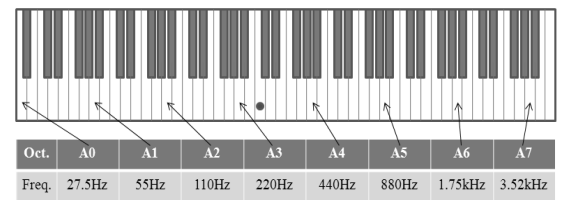


그림 4. 음계에서의 옥타브와 주파수 관계  
Fig. 4. The correlation between octave and frequency in musical scales.

음은 440 Hz의 음높이를, 1초에 880번 진동하는 음은 880 Hz의 음높이를 가진 A음을 낸다. 따라서 주파수가 높은 음은 높게, 주파수가 낮은 음은 낮게 들린다. 그림 3은 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.

음의 높이에서 발견되는 또 하나의 중요한 성질 중 하나가 옥타브이고, 이는 주파수 비가 1대 2인 음정, 즉 완전 8도 음정을 의미한다. 그림 4는 피아노 음계에서의 옥타브와 주파수의 관계를 보여주고 있다. 그림에서 알 수 있는 바와 같이 옥타브사이의 주파수는 배수의 관계를 지니고 있다. 본 연구에서는 가시주파수 영역의 컬러영상 이미지의 특징요소를 추출하여 가청주파수 대역의 특징요소로 각각 변환한 후, 피아노 (88 키보드)의 주파수 대역으로 제한하여 음원을 생성하였다. 입력 컬러영상이 가진 정보를 사운드의 정보로 변환하는 것이므로 음악적인 요소를 고려하여 88 키보드의 주파수 대역으로 제한하였다.

사운드의 두 번째 요소로서, 음색에 영향을 끼치는 하모닉 성분은 음파를 구성하는 고조파의 구성 상태, 즉 고조파 성분의 구조에 따라서 다르게 느껴진다. 예를 들면,

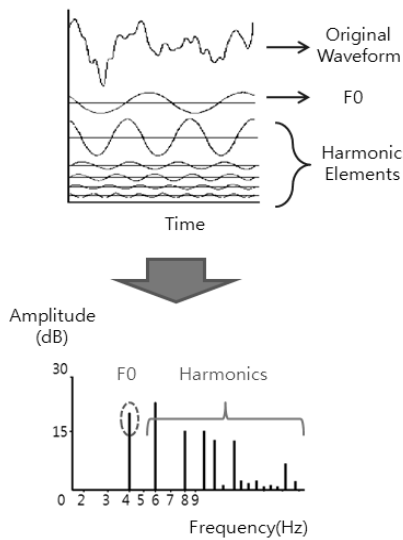


그림 5. 기본주파수와 하모닉 요소  
 Fig. 5. The structure of fundamental frequency (F0) and harmonics.

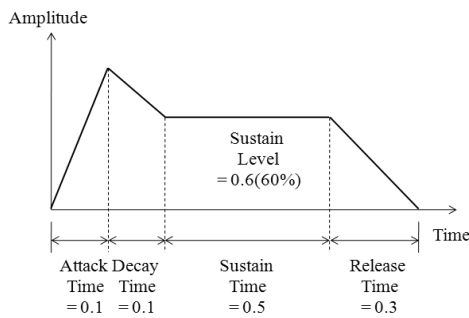


그림 6. ADSR 엔벨롭 신호 생성기의 초기값 설정  
 Fig. 6. Initial setting of ADSR envelop signal generator.

외형상으로 매우 비슷한 악기라고 해도 고조파 배열과 크기가 다르면 두 악기는 서로 다른 음색을 가진다. 그림 5는 원음의 기본주파수와 하모닉 요소들의 예를 보여주고 있다.

사운드의 세 번째 요소로서, 음의 크기를 들 수 있고, 위 그림 5에서 나타낸 주파수 스펙트럼과 함께 사운드의 음색에 영향을 미치는 중요한 요소의 하나가 엔벨롭(envelop)이다. 이는 사운드 발생 후 시간의 경과에 따른 음량변화를 나타낸다. 그림 6은 본 연구의 시뮬레이션 실험에서 사용된 생성 음원의 ADSR (Attack Decay Sustain Release) 엔벨롭 신호의 초기값 설정치를 보여주고 있고, 기준 음량은 80 dB이다.

본 연구에서는 음원 생성을 위한 틀로서, 1985년 M.I.T. 교수인 Barry Vercoe에 의해 개발된 디지털 사운드 합성 프로그램인 Csound를 이용하여 음원을 합성 및 생성하였다. 이를 위하여 우선 입력 RGB기반 컬러이미지를 HSI기

반 컬러모델로 분리한 후, 컬러영상의 세 가지 특징 요소들을 소리의 요소들로 각각 변환한다. 다음 단계로서, ORC (orchestra) 파일과 SCO (score) 파일을 자동 생성하고, Csound의 신호처리 모듈을 이용하여 웨이브(wav) 포맷 음원을 생성하였다.

### III. 컬러이미지-사운드 변환의 제안된 방법

컬러의 세 가지 속성으로 색상, 명도 및 채도로서 표현할 수 있고, 사운드에도 그 음의 특성을 구분할 수 있는 속성으로 음의 높이, 음색 및 음의 크기가 있다. 본 연구에서 컬러이미지-사운드 변환은 컬러이미지의 색상, 명도 및 채도를 소리의 피치, 옥타브 및 하모닉 요소에 각각 대응시키고, 대응 변환된 사운드의 요소를 합성하여 웨이브 포맷 파일로 변환한다. 시뮬레이션 실험에서, 음의 크기는 고정하였고 엔벨롭은 초기 설정치를 이용하였다. 변환 원리로서, 주파수 변환공식을 유추하여 단순 변환 [7-8]시키는 것이 아니라, HSI 컬러모델에 기반한 컬러 요소 각각에 사운드의 세 가지 요소를 대응시키는 방법에 주안점을 두고 있다. 여기서, 사운드의 세 가지 요소를 추출하기 위하여 HSI 컬러 히스토그램을 이용하였다.

그림 7은 HSI 컬러모델 중 색상 성분의 대표적인 RGB 컬러를 음의 높이를 나타내는 음계 성분 중 대표적인 화음인 도미솔로의 대응 과정을 보여주고 있다. 일반적으로 가시영역의 대표적인 RGB 주파수 비율은 음계의 도미솔 성분의 주파수 비율과 1:4/5:2/3의 비율로 일치한다고 알려져 있다. 이 원리와 함께, HSI 컬러모델의 색상 성분은 RGB 상호간에 120도의 관계를 가진다는 정보를 바탕으로, 본 연구에서는 색상 성분을 30도씩 분리하여 피아노의 12음계로 대응시켰다. 즉, 기존연구 [7-8]에서처럼 주파수 변환 공식에 따라 12음계의 주파수비와 일치하도록 색상과 음을 대응시키지 않고, 색상을 동일 간격이 되도록 30도씩 그룹화하여 12음계로 대응시키는 비교적 단순하고 직관적인 방법을 이용하였다.

그림 8은 컬러이미지에서 사운드 변환의 전체적인 순서도를 보여주고 있다. 입력 RGB 컬러모델 기반 입력영상은 우선 HSI 컬러모델로 변환되고, 이 중 색상 성분의 각각의 픽셀열에서 히스토그램을 구하고, 그 중 최대값을 가진 성분을 12평균을 중 대표 음계로 정한다. 같은 방식으로 명도 성분의 히스토그램을 구하고, 그 중 최대값을 가진 성분을 7 옥타브 중 대표 옥타브로 정한 후, 음계와 옥타브 정보를 기반으로 기본주파수를 결정한다.

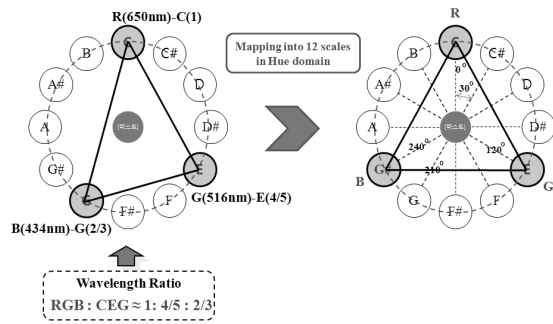


그림 7. 가시주파수에서 음계 성분으로의 변환  
 Fig. 7. The conversion of visible frequencies into musical scale elements.

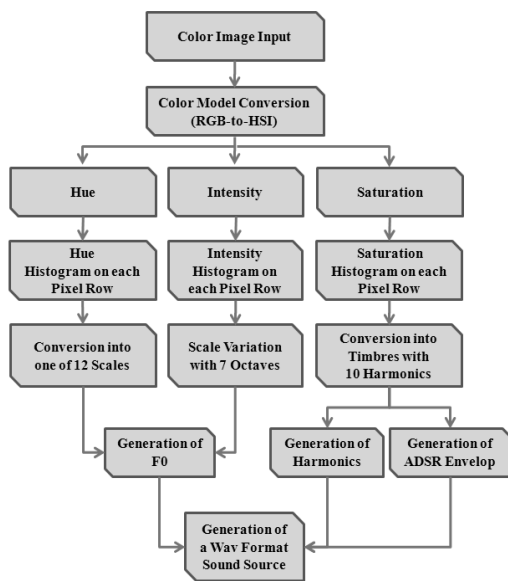


그림 8. 컬러이미지에서 사운드 변환의 순서도  
 Fig. 8. The flow diagram of the conversion of color image into sound.

또한, 채도 성분의 히스토그램을 구하고 총 10개의 하모닉 성분을 정하고 ADSR 엔벨롭 성분도 결정한다. 최종적으로, 모든 정보를 통합하여 음원을 생성한다.

그림 9는 히스토그램을 이용한 컬러이미지-사운드 변환의 전 과정의 개념도를 간략히 보여주고 있다. 임의의 컬러영상이 입력되면, 주어진 영상을 화소 단위로 분석하여 RGB 정보를 추출한다. 추출한 RGB 정보는 HSI와 같이 색상, 채도, 명도를 기준으로 하는 좌표 체계의 정보로 바꾼다. 본 연구에서는 피아노 (88키보드) 음역을 대상으로 12평균율과 7개의 옥타브를 기반으로 구성하였다. 음색은 영상의 채도성분을 기반으로 10개의 하모닉 성분을 가진 음색으로 구성하였다. 그림 9의 (a)는 입력 컬러이미지로부터 색상과 명도 히스토그램으로부터 음계와 옥타브 결정 과정의 예를 보여주고 있다. 변환된

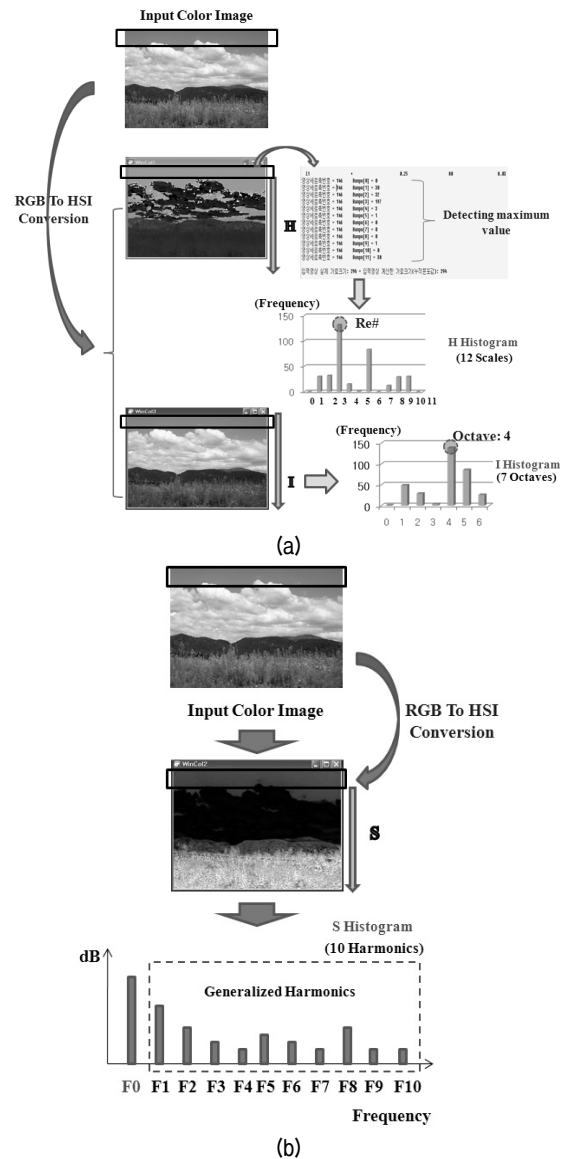


그림 9. HSI 컬러 히스토그램을 이용한 색-음 변환 (a) 입력이미지로부터 색상과 명도 히스토그램으로의 생성과정 (b) 입력이미지로부터 채도 히스토그램으로의 생성과정  
 Fig. 9. Color image into sound conversion using HSI color histograms. (a) The generation process from input image to hue and intensity histograms (b) The generation process from input image to saturation histograms.

HSI 값 중 색상 채널의 각각의 가로축 픽셀 값들을 기준으로 12개의 그룹으로 나누어 히스토그램을 구하고, 히스토그램의 최대값을 12평균율의 하나로 대응시킨다. 그림의 예는 4 옥타브의 Re#이 최종 결정됨을 보여주고 있다. 그림 9의 (b)는 입력 컬러이미지로부터 채도 히스토그램을 바탕으로 하모닉 성분의 생성을 보여주고 있다. 채도 채널의 가로축의 픽셀 값들은 위와 동일한 방법으로 10개의 그룹으로 나누어 히스토그램을 만들고, 음색에

영향을 끼치는 하모닉 주파수 성분에 대응시켰다.

### IV. 실험 및 결과

제안된 컬러이미지-사운드 변환 원리를 구현하기 위한 도구로서 Microsoft Visual C++과 Csound를 사용하고, 생성된 음원의 평가 도구로서 Praat를 사용하였다. 즉, RGB 컬러영상에서 HSI 컬러모델로의 변환과 컬러이미지-사운드 변환 프로그램을 C++로 코딩하였고, 음원 생성을 위해 Csound를 사용하였다. 또한, 생성된 음원의 분석 툴로는 Praat을 이용하여 분석하였다. 그림 10은 HSI 히스토그램을 이용한 컬러이미지-사운드 변환 과정을 거쳐 변환, 생성된 음향파일을 Praat를 이용하여 스펙

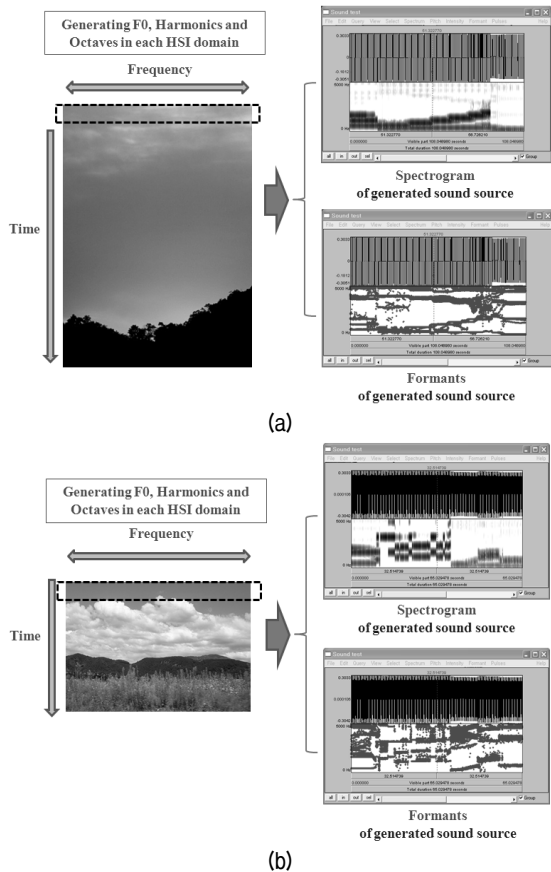


그림 10. 생성된 음원들의 스펙트로그램과 포먼트 분석 (a) 입력 컬러영상 '노을(321\*432 pixels)'로부터 생성된 음원의 스펙트로그램과 포먼트 분석 (b) 컬러영상 '들판(387\*260 pixels)'의 스펙트로그램과 포먼트 분석  
 Fig. 10. The analysis of both spectrograms and formants of the generated sound sources. (a) The analysis of both spectrogram and formant of the generated sound source from input color image 'sunset' (b) The analysis of both spectrogram and formant of the generated sound source from input color image 'green field'.

트로그램 (Spectrogram) 및 포먼트 (Formants)를 추출한 결과를 보여주고 있다. 즉, 그림 10의 (a)는 321\*432 픽셀의 컬러영상 '노을'로부터 생성된 음원, 그리고 그림 10의 (b)는 387\*260 픽셀의 컬러영상 '들판'으로부터 생성된 음원의 스펙트로그램과 포먼트 분석 결과를 보여주고 있다. 여기서 보여 지는 바와 같이, 생성 음원의 스펙트로그램 및 포먼트 성분 추출에서 각각 기본주파수 및 하모닉 성분을 확인할 수 있었다.

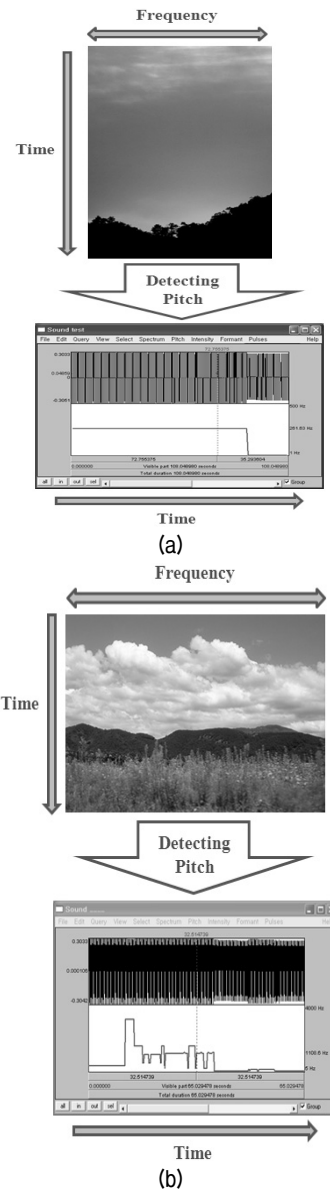


그림 11. 생성된 음원의 피치 분석 (a) 입력 컬러영상 '노을'로부터 생성된 음원의 피치 분석 (b) 입력 컬러영상 '들판'로부터 생성된 음원의 피치 분석  
 Fig. 11. The analysis of the pitch of generated sound source. (a) The analysis of the pitch of the generated sound source from input color image 'sunset' (b) The analysis of the pitch of the generated sound source from input color image 'green field'.

본 연구에서는 RGB 컬러모델에서 HSI 컬러모델로 변환한 후, H, S 및 I 각각의 채널의 세로 픽셀영역을 시간 축으로 계산하였다. 그리고 색상 채널의 가로 픽셀영역은 히스토그램을 이용하여 최대값을 가진 성분을 12평균을 중 대표 음에 해당하는 값으로 매핑하고, 명도 채널의 가로 픽셀영역에서 같은 방식으로 최대값을 가진 성분을 구한 후 옥타브에 해당하는 값으로 매핑하여 기본적인 멜로디 성분을 합성하였다.

그림 11은 두 가지 상이한 컬러이미지들을 입력받아 생성된 음원들로부터 Praat를 이용하여 피치를 분석한 결과를 보여주고 있다. 그림 11의 (a)에서는 대부분을 차지하는 하늘의 배경 영역과 산의 이미지 영역에서 색상과 명도가 크게 변함으로 피치값도 그에 따라 변동됨을 알 수 있었다. 그림 11의 (b)에서는 크게 하늘, 구름 및 들판 영역으로 나뉘어 피치값도 그에 따라 변동됨을 알 수 있었다. 이와 같이 색상과 명도 채널로부터 최종 피치값을 추출하고, 이 값들은 입력되는 컬러 이미지의 색상과 명도의 히스토그램에서 구해지는 최대치의 성분 변화에 따라 함께 변화함을 알 수 있었다. 즉, 비교적 비슷한 색상과 명도 영역은 거의 일정한 피치값을 유지함을 볼 수 있었다.

## V. 결론

본 연구는 컬러이미지에서 사운드를 연상하는 인간의 공감각적 기능을 모방하는 지능로봇의 개발을 궁극적인 목표로 하고 있고, 이를 위하여 컬러이미지-사운드 상호 변환에 기초한 응용 시스템의 구현을 부차적인 목표로 하고 있다. 그 첫 번째 기초연구로서, 컬러 이미지를 사운드로 변환하는 방법론을 바탕으로 기본 시스템을 구현하였다. 시뮬레이션 결과, 입력 RGB 컬러 영상을 HSI 영상으로 변환 분리하여 각각의 색상, 채도 및 명도 히스토그램을 구하고, 이들로부터 음의 높이, 하모닉 및 옥타브에 각각 대응시킴으로써 최종 음원이 생성됨을 확인하였다. 그러나, 본 연구의 컬러이미지-사운드 변환을 바탕으로 생성된 음원의 사운드는 뚜렷하게 출력 되지만 음악적인 요소가 배제되어 있어 음원만 듣기에는 부족한 점이 있었다.

앞으로의 연구 방향은 컬러이미지-사운드 변환관계에 있어 컬러영상의 감성정보를 추출하여 음악적인 요소정보로 대응시키는 작업이 추가 되어야 할 것이다. 이를 위해 영상의 3요소인 색상, 질감 및 형태의 정보추출이 필

요하고, 추출된 영상의 3요소를 음악의 3요소인 화음, 멜로디 및 하모니에 대응시키는 최적의 방식을 찾아야 하고, 동시에 인간의 공감각적 기능이 얼마나 유효하게 추출, 적용되었는지를 실제 보고 듣는 인지실험을 통하여 규명되어야 할 것이다.

## 감사의 글

이 논문은 2010년도 정부 (교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (2010-0015818).

## 참고 문헌

1. 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.
2. 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.
3. 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.
4. Peter B.L. Meijer, "An experimental system for auditory image representations", *IEEE trans. on bio-medical engineering*, vol. 39, no. 2, pp. 112-121, 1992.
5. 이형민, *영상악기 만들기 1*, 마이크로소프트웨어, 10월호, pp. 234-239, 2006.
6. 손우람, *영상악기 만들기 2*, 마이크로소프트웨어, 11월호, pp. 232-238, 2006.
7. 김길호, 백정기, *사운드컬러하모니즘: 음악이 흐르는 컬러 배색 사전*, 임프레스, 2003.
8. 김길호, *화성법을 이용하여 색과 음을 상호변환하고 색채를 조화하는 방법 및 장치*, 출원번호 10-99-34242, 1999.
9. 김길호, *화성법을 이용한 음색변환표*, 출원번호 10-2001-0087651, 2001.
10. 김만호, 정성환, 강민수, "음의 시각화와 그 표현의 경향", *디자인/學研究*, 16권, 3호, 101-110쪽, 2003.
11. 박진희, "음악과 색채의 유기적 관계에 관한 연구: 모드를 중심으로", *상명대 뉴미디어정보통신대학원 석사논문*, 2005.
12. Michael Freeman, *Mastering Color Digital Photography*, Ilex Press, 2004.
13. 김유라, "소리의 표현 연구", *서울여자대학교 대학원 석사학위논문*, 1998.

## 저자 약력

• 김 성 일 (Sung-Il Kim)

한국음향학회지 제23권 제3E호 참조