

## 증강 현실 기반의 실시간 음악 편집 및 재생 시스템

김은영\*, 오동열\*\*

### A Realtime Music Editing and Playback System in An Augmented Reality Environments

Eun-Young Kim\*, Dong-Yeol Oh\*\*

#### 요약

본 논문에서는 증강 현실에서 객체 구분을 위하여 사용되는 다양한 마커를 기반으로 실시간으로 음악을 편집하고 이를 재생할 수 있는 시스템을 설계 및 구현한다. 제안 시스템은 증강 현실 마커 기반의 뮤직 마커와 뮤직 보드로 구성된다. 뮤직 보드상에 배치된 뮤직 마커의 내용을 악기의 종류와 미리 정의된 재생 트랙과 매핑하고 2차원 평면상에서 상대적인 뮤직 마커의 위치 값을 측정하여 이를 midi 트랙에서 공간 효과와 관련있는 파라미터 값으로 설정한다. 제안된 시스템은 최대 1600×1200 픽셀을 지원하는 캠을 기반으로 다양한 해상도에서 마커의 움직임이나 추가에 따른 재생 지터 시간을 측정하여 성능을 평가하였으며, 실험 결과 860×600 픽셀에서 초당 두 개의 프레임 을 처리하는 경우, 마커의 인식률이 높았고 허용 가능한 지터 내에서 안정적인 결과를 도출하였다. 제안 시스템은 오늘날 대체 의학으로 제시되고 있는 음악을 기반으로 한 감성 치료법이나 음악적 배경이 없는 어린이나 노약자도 손쉽게 사용할 수 있어서 교육적인 측면에서도 활용될 수 있다.

▶ Keyword : 증강현실, 휴먼 컴퓨터 인터랙션, 음악 재생 시스템, midi

#### Abstract

In this paper, We propose real-time sound editing and playback systems which is based on Augmented Reality. The proposed system are composed with music maker which is based on AR maker and music board. By using music marker's contents, the proposed system selects the kinds of musical instruments and pre-defined midi track and by calculating the relative location of music marker on 2-dimensional plane, we set the spatial relative parameter in midi track. For performance evaluation, we check the jitter value of in various resolutions by using CAM which supports 1600×1200 as the maximum resolution. As a result, when we set the configuration value

• 제1저자 : 김은영 • 교신저자 : 오동열  
• 투고일 : 2011. 02. 21, 심사일 : 2011. 03. 01, 게재확정일 : 2011. 03. 14  
\* 안산공과대학 멀티미디어 콘텐츠과(Dept. of Multimedia Contents, Ansan Colleague)  
\* (주)인젠트 연구개발부(R&D Division, Inzent Cooperation)  
※ 이 논문은 2010학년도 안산공과대학 학술연구비에 의하여 연구된 것임

of CAM as 860×600 pixels and process two frames per minute, the success ratio of recognizing music markers and jitter values are acceptable. It can be utilized in the fields of alternative c machine which is based on music and also be utilized in the educational aspects because child or elderly who don't know enough musical theory can easily handle it.

▶ Keyword : Augmented Reality, HCI, Music Playback System, MIDI

## I. 서 론

음악은 오랜 기간 동안 사람의 심리적 안정과 즐거움을 제공하기 위하여 다양한 형태와 기술로 발전되어 왔다. 아날로그 시대에는 악기 연주의 녹음과 음악의 편집뿐만 아니라 음악의 재생도 모두 아날로그 기반 매체에 저장되어 있고 저장된 매체는 특정 기기를 이용하여 재생하게 되는데, 이러한 악기의 녹음과 음악의 편집 및 재생에 관련된 일련의 과정은 전문적인 교육과 기술 없이 일반인들이 접근하기 어려운 분야였다. 아날로그 기반의 매체가 컴퓨터의 보급과 더불어 디지털 매체로 바뀌면서 미디(MIDI : Musical Instrument Digital Interface) 신호를 기반으로 한 신디사이저와 같은 전자 악기의 연주 및 편집과 더불어 컴퓨터 기반의 레코딩이 대중화되면서 더 이상 음악의 편집과 재생은 특정인들만의 소유물이 아니다[1].

그러나 이러한 컴퓨터의 발전에도 불구하고, 전자악기의 습득과 더불어 음악 작곡 및 편곡을 위한 시퀀싱 프로그램에 관한 지식 없이는 음악의 편곡은 여전히 일반인들에게 어려운 대상이며, 대부분의 사람들은 디지털 매체로 바뀐 MP3 플레이어와 같은 재생기를 통한 음악 감상이 대부분을 이루고 있다. 이에 본 논문에서는 증강 현실 환경에서 객체 구분을 위해서 사용되는 다양한 마커를 확장하여 마커의 종류와 위치에 따라서 미디 포맷으로 미리 저장된 해당 트랙을 반복적으로 재생함으로써 음악을 구성하고 있는 악기와 멜로디를 실시간으로 편곡하고 편곡된 곡을 재생할 수 있는 증강 현실 기반의 실시간 음악 편집 및 재생 시스템을 제안한다. 제안 시스템은 특별한 음악에 대한 배경 지식과 시퀀싱 프로그래밍에 대한 교육 없이 직관적인 방식으로 선택된 미디 트랙을 재생할 수 있으며 마커의 매핑을 전방향 추론이 가능한 규칙 기반으로 제공함으로써 보다 다양한 형태의 악기 설정이나 트랙 설정이 가능하다. 제안 시스템은 오늘날 또 다른 의미로 제시되고 있는 음악을 기반의 감성 치료법이나 음악적 배경이 없는 어린이나 노약자도 손쉽게 사용할 수 있어서 교육적인 측면에서도 활용될 수 있다.

## II. 관련 연구

2장에서는 증강 현실을 기반으로 악기의 연주나 음악의 재생과 관련된 기존에 관련 연구를 살펴보고, 각 관련연구에서 제시하는 시스템들 간의 특징과 장단점에 대해서 설명한다.

### 1. 기존 관련 연구

#### 1.1 Augmented Groove 기반의 음악 필더 적용

Poupyrev는 2000년에 [그림 1]과 같이 Augmented Groove라는 음악 재생을 위한 컨트롤러를 제안하였다[2].

제안된 Augmented Groove는 레코드판 형태에 마커를 장착한 가상의 컨트롤러로 사용자는 Augmented Groove의 상하 위치와 회전 및 기울기를 변형하는 방식으로 재생 중인 미디 음악의 반복구간에 다양한 형태의 필터를 적용할 수 있는 미디 재생 기반의 시스템을 제시하였다. 제안하는 Augmented Groove는 마커의 높이와 기울기 및 회전 방향을 인식하여 이를 기반으로 재생되는 음악에 다양한 효과를 적용하였다. 본 연구는 증강 현실을 기반으로 직관적인 인터페이스를 제시하였으나 적용 가능한 범위를 음악 재생에 관련된 필터에 국한하였다는 단점이 있다.

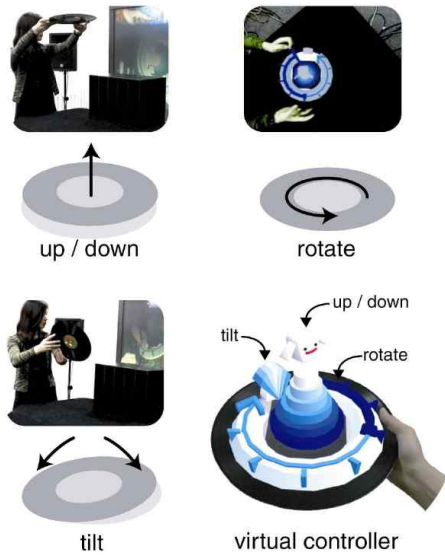


그림 1. Augment Groove기반의 음악적 인터페이스 제스처  
Fig 1. Gestural Musical Interface in Augmented Groove

1.2 탠저블 인터페이스 기반의 작곡 및 재생

탠저블 인터페이스(Tangible Interface)는 가상 공간에서 현실과 같은 느낌을 경험할 수 있도록 실감성을 부여하는 기술로 Costanza는 2003년도 연구에서 다양한 마커를 이용하여 음악의 재생과 편집을 가능케 하였다[3]. 제안 시스템에서는 다음과 같이 세 가지 적용 가능한 분야를 제안하였다. 첫째는 악보 모양과 마커를 이용한 오선 배치를 통해 멜로디를 작성할 수 있는 방법이며 둘째로는 루프 기반의 드럼 패턴을 재생할 수 있는 방법을 제안하였고 마지막으로 루프 재생 및 이펙터를 적용할 수 있는 방식을 제안하였다.

[그림 2]는 탠저블 인터페이스를 기반으로 멜로디 작성, 드럼 패턴 재생 및 루프 재생과 이펙터 적용의 예를 나타내고 있다.



그림 2-1. 멜로디 작성 예  
Fig 2-1. Example of Composing a Melody

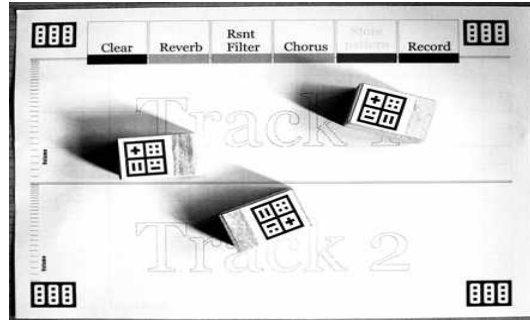


그림 2-2. 드럼 패턴 재생 예  
Fig 2-2. Example of Playing Drum Patterns

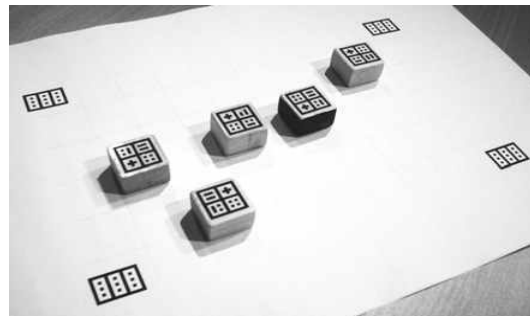


그림 2-3. 루프 재생 및 이펙터 적용  
Fig 2-3. Playback Loop and Apply Effector

1.3 MMM 기반의 모션과 효과음의 결합

2004년도에 KIA는 카메라를 기반으로 무용가의 신체적인 움직임을 인식하고 이에 맞추어 실시간으로 연관된 음악이나 효과음을 재생 할 수 있는 시스템을 제안하였다. 관련 연구에서는 그림 3과 같이 MvM(Music Via Motion)이라는 무용가의 신체적인 움직임이나 얼굴 인식의 결과에 맞추어 실시간으로 음향이나 음악을 재생할 수 있는 프레임워크를 제안하였다. 해당 연구에서는 마커없이 사람의 행위를 인식하는 과정에서 제한된 환경과 미인식율에 따른 음악 재생의 매핑 오류라는 단점이 있다[4].

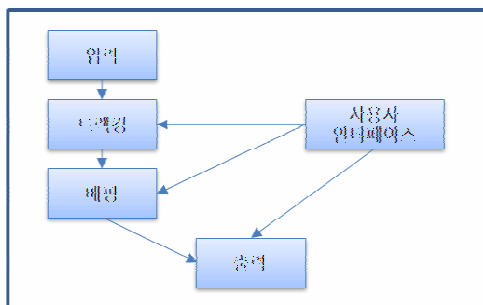


그림 3. MMM의 주요 모듈  
Fig 3. Main Modules of MMM

### 1.4 기타 학습을 위한 증강 현실 연구

2005년도에 Liarokapis는 [그림 4]와 같이 여러 개의 마커에 매핑된 기타 코드의 운지법을 디스플레이 함으로서 보다 손쉽게 기타를 학습할 수 있는 인터페이스를 제시하였다[5].

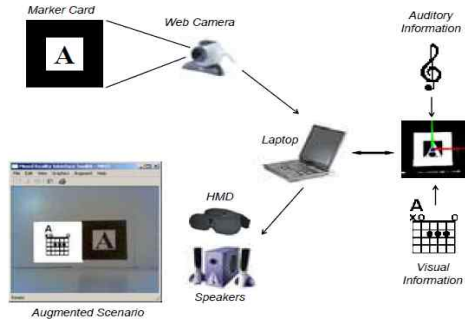


그림 4. 관련 연구 시스템 개요

Fig 4. System Overview of Relative Study

관련 연구에서는 기타의 운지법 뿐만 아니라 해당 코드를 재생함으로써 시각적인 효과와 청각적인 효과를 제공하고 있다.



그림 5. 마커 인식과 가상 손 모델

Fig 5. Marker Detection and Virtual Hand Model

유사한 연구로 Motokawa는 기타의 넥 부분에 마커를 위치시켜서 해당 코드에 따른 운지법을 그래픽으로 변환된 가상의 손을 이용하여 실사에 매핑함으로써 보다 손쉽게 해당하는 코드의 운지가 가능한 방식을 [그림 5]와 같이 제공하였다[6].

### 1.5 가상 공간상에 악기 연주에 관한 연구

Smith는 2007년도에 마커가 위치된 공간상에 가상 하프를 실시간으로 렌더링하고, 교차된 사람의 손이나 사물의 움직임에 따라서 가상의 하프를 연주할 수 있는 가상 하프(Virtual Harp) 시스템을 제안하였다[7]. 관련 연구에서는 실제 하프의 연주와 동일한 효과를 위해서 움직임에 따른 하프 줄의 진동 모델을 제안하였다. 제안하는 진동 모델은 캠에서 사람의 손에 의해서 가려지는 영역의 시작과 끝을 인식하

여 음의 높이와 진동의 정도를 시뮬레이션하여 미디어 기반의 전자악기에서 하프 소리를 재생하는 시스템으로 음악을 기반으로 한 대한 치료로 활용 가능하다. [그림 6]은 마커에 도식화된 가상 하프의 예를 나타내고 있다.

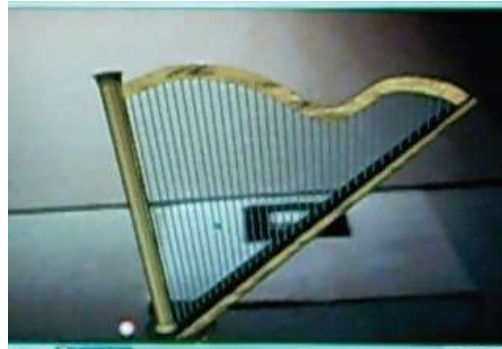


그림 6. 증강 현실에서의 가상 하프

Fig. 6. Virtual Harp in Augmented Environment

## 2. 기존 연구 비교 분석

증강 현실과 음악과 관련된 기존 연구는 KIA가 제공한 MvM 프레임워크와 비교하여 다음과 같이 크게 5가지 측면에서 분석 하였다.

첫째로 시스템과 사용자 간의 협업 방식(입력, 협업에 따른 인지 방식(트래킹), 셋째로는 트래킹된 입력 데이터와 실제 서비스와의 연관관계(매핑), 넷째로 제공되는 서비스(출력)로 크게 구분하였으며 마지막으로는 사용자 인터페이스에 따른 차이점을 구분하였으며 관련 연구별 비교 결과 [표 1]과 같다. Poupyrev 연구의 경우, 재생 중인 음악에 필터를 적용할 수 있는 방안을 제시하였으나, 적용 가능한 대상이 음색을 변화하기 위한 필터로 국한되어 있다. Costanza의 경우, 다양한 마커를 이용하여 멜로디 작성, 드럼 패턴 재생 및 루프 재생과 이펙터 적용이 가능하나, 사용된 마커 정보만 가지고 그 의미를 명확하게 전달할 수 있는 부분이 상대적으로 부족하다. KIA의 연구는 MvM을 이용하여 시스템을 효과적으로 구성할 수 있는 장점이 있으나, 마커가 없는 증강 현실 모델을 적용함으로써 획득된 영상의 오인식에 의하여 사용자가 의도하지 않은 결과가 도출될 수 있는 불확실성이 포함되어 있다. Liarokapis와 Motokawa의 연구는 음악 재생이 아닌 기타 학습의 협업 도구로서 새로운 방향을 제시하였고, Smith는 연구에서 가상의 하프를 연주할 수 있는 기법을 제시하였으나 기본적인 하프 악기의 특성을 알아야만 자신이 원하는 연주를 할 수 있는 전문적인 학습이 필요하다는 단점이 있다.

표 1. 관련 연구 비교

Table 1. Comparison of Relative Research

관련 연구	입력 대상	트래킹	매핑	출력	UI
Poupyrev 연구	마커	상하위치 회전 기울기	필터	필터가 적용된 사운드	Augmented Groove
Costanza 연구	마커	마커 위치 마커 내용	멜로디 드럼패턴 루프 재생	마커에 따른 멜로디 마커에 따른 드럼 패턴 마커에 따른 루프 재생	다중 마커 배치
KIA 연구	사람의 움직임	사람의 움직임 안면 인식	효과음 멜로디	미리 정의된 사운드	캠 영상 마커
Liarokapis 연구	마커	마커 내용	코드 마커	기타 운지법 해당 코드 사운드	코드 마커
Motokawa 연구	마커	마커 내용 마커 위치	코드 마커	가상 운지 모델(손)	코드 마커
Smith 연구	마커	사람의 손	해당 스트링	가상 악기 재생	마커

### III. 증강 현실 기반의 음악 편집 및 재생 시스템의 설계

#### 1. 뮤직 보드와 뮤직 마커의 모델링

[그림7]과 같이 제안 연구에서는 뮤직보드(Music Board) 라는 평면형태의 공간을 정의하고 뮤직 보드상에 배치하는 각각의 증강 현실 마커를 뮤직 마커(Music Marker)로 제안한다.

뮤직보드는 실제 사용자가 무대에서 진행되는 연주를 감상한다라는 가정 하에 연주가 진행되는 무대를 의미하는데, 실제 시스템에서는 수직으로 투영하여 캠을 통해서 획득된 영상의 배경이 뮤직보드에 해당한다. 뮤직마커는 마커 각각이 특정 미디 트랙을 반복하여 설정된 악기를 기반으로 연주되는 정보를 나타내는데 이는 캠을 통해서 획득된 마커 영상의 2차원 평면상에서의 위치 정보와 마커에 출력된 식별자 정보를 의미한다.



그림 7. 뮤직 보드와 뮤직 마커의 정의  
Fig 7. Define Music Board and Music Marker

디지털 악기와 컴퓨터간에 표준 통신 규약으로 제시된 미디포맷에서는 표준 컨트롤을 이용하여 해당 미디 트랙의 재생 시 스테레오에서 좌우 설정이 가능하고 이와 더불어 해당 트랙의 볼륨 설정이 가능하다[8]. 이에 본 논문에서는 공간각적인 소리를 현실 적으로 증강하기 위해서 재생되는 미디 트랙의 좌우 설정(Midi Control Number 10 - Pan)과 볼륨 설정(MIDI Control Number 7 - Volume)을 뮤직 보드 2차원 평면상의 뮤직 마커의 위치에 따라서 변경한다. (A)는 마커 위치에 따른 미디 컨트롤의 Pan 값을, (B)는 미디 컨트롤의 Volume 값을 계산하는 알고리즘을 나타낸다.

$$Pan = MM_x / MB_{width} \times P_{Max} \dots\dots\dots (A)$$

$$Volume = (MB_{Height} - MM_y) / MB_{Height} \times (V_{Max} - V_{Min}) + V_{Min} \dots\dots\dots (B)$$

제안하는 (A)와 (B)에서 MM<sub>x</sub>와 MM<sub>y</sub>는 인식된 뮤직마커의 X, Y좌표를 의미하고 MB<sub>width</sub>와 MB<sub>height</sub>인식된 뮤직 보드의 너비와 높이를 나타낸다. P<sub>min</sub>과 P<sub>max</sub>는 Pan 컨트롤의 최소값과 최대값을 나타내고, V<sub>min</sub>과 V<sub>max</sub>는 Volume 컨트롤의 최소 값과 최대 값을 의미한다. 본 수식에서 좌표값과 공간 영역에 대한 크기를 나타내는 수치 값은 영상을 통해서 획득된 픽셀 값을 의미하며, Pan 컨트롤 설정을 위한 값은 0~127까지의 값을 사용하고 Volume 컨트롤 설정의 경우에는 87~127까지의 값을 사용한다. 컨트롤의 설정 범위의 임계치는 임의로 설정하여 적용할 수 있으나 MIDI 규약상, 설정값은 0~127의 범위 내에서 값으로 설정한다.

#### 2. 뮤직 마커의 생성과 캠을 이용한 인식

미디 파일에서의 미디 트랙과 해당 트랙의 악기를 나타내는 정보는 [표 2]와 같이 영문 대문자와 숫자로 조합된 마커를 뮤직 마커로 사용한다. 제안하는 뮤직 마커의 영문 대문자는 재생할 미디 트랙의 악기의 종류를 의미하며 숫자는 멜로디 패턴에 부여된 시퀀스 정보를 나타낸다.

표 2 악기와 패턴에 따른 마커의 종류  
Table 2 AR marker by Instrument and Pattern Info

패턴	Guitar	2nd Guitar	Base	Piano	Drum
1	G 1	S 1	B 1	P 1	D 1
2	G 2	S 2	B 2	P 2	D 2
3	G 3	S 3	B 3	P 3	D 3
4	G 4	S 4	B 4	P 4	D 4
5	G 5	S 5	B 5	P 5	D 5
6	G 6	S 6	B 6	P 6	D 6
7	G 7	S 7	B 7	P 7	D 7

생성된 뮤직 마커의 인식을 위해 그림 8과 같이 Marker Generator를 이용하여 16 X 16의 세그먼트를 기반으로 각 뮤직 마커에 해당하는 패턴을 생성한다. 제안 연구에서는 다음과 같이 총 5개의 악기와 각 악기별로 7개의 멜로디 패턴을 사용하기 때문에 16 X 16 세그먼트를 기반으로 한 패턴으로도 35개의 미디 트랙을 조합하여 표현 할 수 있다.

캠을 통해서 획득된 영상을 인식하기 위해서 FLARToolkit을 사용한다. FLARToolkit을 통해서 획득된 정보는 뮤직 마커의 종류와 뮤직 마커의 3차원 공간 좌표 (X, Y, Z)가 되는데 제안하는 시스템에서는 2차원 평면상의 정보인 (X, Y) 좌표와 뮤직 마커의 정보만을 이용한다. [그림 8]은 뮤직 마커를 기반으로 패턴 파일을 생성한 예를 나타낸다.

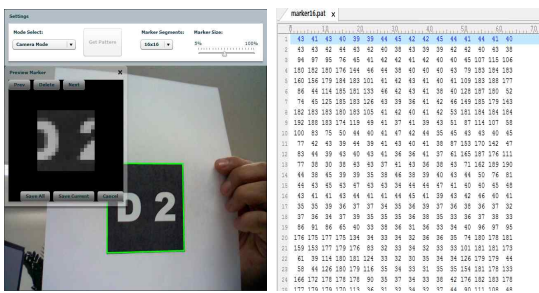


그림 8. 드럼 패턴 2 뮤직 마커 생성과 패턴 파일의 생성 예  
Fig 8. Example of Drum Pattern 2 Music Marker and Pattern File

### 3. 뮤직 마커와 재생 트랙 정보 모델링

캠을 이용하여 획득된 뮤직 마커 정보를 기반으로 미디 파일 내 정의된 트랙들 중에서 재생할 트랙을 선택하기 위한 규칙은 김은영이 기존 연구에서 제시한 상황 인식 미들웨어의 서비스 추론을 위한 일차 술어 논리(First-Order Predicate Logic)를 확장하여 정의한다[9]. 뮤직 마커를 인식한 결과는 다음 (C)와 같은 형태로 해당 미디 트랙의 검색을 위하여 파라미터로 표현되고 전달된다.

$MM(Inst = a, Pattern = b, X = c, Y = d)$   
*Inst*: 해당 악기(인식된 뮤직 마커의 영문 대문자)  
*Pattern*: 해당패턴이 저장된 트랙(인식된 뮤직 마커의 숫자)  
*X*: 인식된 뮤직 마커의 X좌표  
*Y*: 인식된 뮤직 마커의 Y좌표

..... (C)

표 3. 뮤직 마커 정보간에 관계  
Table 3. Relations between Information Sets

상황 정보	정의
$A \wedge B$	상황 종류가 A이거나 B인 경우
$A \vee B$	상황 종류가 A이고 B인 경우
IF A THEN B	상황 종류 A 이면 B를 실행서비스추론
$p = v$	IF 절에서는 변수 p의 값이 v 임을 의미 Then 절에서는 변수 p의 값을 v로 설정

캠을 통해서 획득된 뮤직 보드상에 존재하는 다수의 뮤직 마커 정보는 상황 정보를 나타내는 관계 연산자로 기술되며 기존 연구에서 제안한 관계 연산자를 [표 3]과 같이 확장한다.

뮤직 마커 정보를 획득하게 되면 미리 제작된 미디 파일에서 해당 트랙에 대한 검색과 마커의 정보 및 좌표에 따른 컨트롤 값을 설정하게 되는데, 이를 (D)와 같이 정의한다.

$Track(Inst = a, TrackNum = b, Pan = c, Volume = d)$   
*Inst*: 설정할 악기 번호  
*TrackNum*: 해당패턴이 저장된 트랙 번호  
*Pan*: Pan 설정 값  
*Volume*: Volume 설정 값

..... (D)

(E)는 획득된 마커 정보들을 기반으로 해당 트랙에 대한 컨트롤 값을 설정과정에서 인식된 뮤직 마커가 B2와 D3인 경우에 해당 컨트롤을 정방향 추론(Forward Chaining)을 통해서 설정하는 예로 B2 마커의 경우, 재생할 트랙을 12번으로 설정하고 악기는 GM(General Midi)의 33번에 해당하는 Acoustic Base로 설정하고 Pan 값은 16, 볼륨 값은 117로 설정하라는 것을 의미한다[10].

$IFMM(Inst = B, Pattern = 2, X = 80, Y = 120)$   
 THEN  
 $Track(Inst = 33, TrackNum = 12, Pan = 16, Volume = 117)$   
 $ANDIFMM(Inst = D, Pattern = 3, X = 320, Y = 240)$   
 THEN  
 $Track(Inst = 1, TrackNum = 30, Pan = 64, Volume = 107)$

..... (E)

### 4. 제안 시스템 설계

본 장에서는 증강 현실 마커를 기반으로 실시간으로 음악을 편집하고 이를 재생할 수 있는 시스템을 설계한다. 제안 시스템은 [그림 9]와 같이 캠을 이용하여 마커를 인식하기 위

한 마커 인식부, 획득된 마커 정보를 가지고 해당 음악의 반복 구간을 검색하기 위한 규칙 엔진 기반의 규칙 검색부와 조회된 정보를 가지고 미디포맷으로 작성된 음악의 반복 구간을 재생하기 위한 미디어 재생부 세 가지로 구성된다.

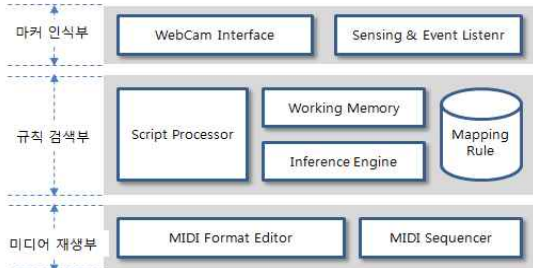


그림 9. 제안 시스템 SW 구성도  
Fig 9. Software Architecture of the Proposed System

첫째, 마커 인식부에서는 웹캠을 이용하여 미리 생성된 뮤직마커를 인식하고 뮤직 마커가 뮤직 보드 상의 2차원 평면 정보와 뮤직 마커의 정보를 생성하고 이를 규칙 검색부에 전달한다.

둘째, 마커 인식부를 통해서 획득된 뮤직 마커 정보를 미리 정의된 해당 미디 포맷 파일의 트랙을 검색하기 위한 규칙 검색부는 다음과 같이 크게 4가지 구성요소로 이루어져 있다.

- 스크립트 처리기(Script Processor) : 정의된 스크립트를 분석하여 명시된 규약에 따라 수행한다.
- 실행 메모리(Working Memory) : 스크립트 처리기를 통해서 획득된 뮤직 마커 정보의 조건 절에 해당하는 규칙을 생성하고 이를 추론 엔진에 전달한다.
- 추론 엔진(Inference Engine) : 정의된 매핑룰로부터 획득된 뮤직 마커 정보가 조건 절을 만족하는 경우, 해당 작업을 실행하거나, 뮤직 마커 정보를 유추한다.
- 매핑 룰(Mapping Rule) : 뮤직 마커 정보와 실제 미디 파일내의 해당되는 트랙에 대한 메타 정보

셋째, 미디어 재생부는 획득된 설정 규칙을 이용하여 해당하는 미디 트랙의 무음/유음을 여부를 설정하고 뮤직 마커를 통해서 획득된 정보를 기반으로 Volume과 Pan을 설정하고 해당 파일을 시퀀서 프로그램을 기반으로 재생한다.

IV. 제안 시스템의 구현 및 성능 테스트

4.1 제안 시스템 구현

4.1.1 제안 시스템 하드웨어 구성

제안 시스템의 하드웨어 구성은 [그림 10]과 같다. 뮤직보드와 뮤직 마커 영상을 촬영하기 위한 웹캠으로 로직텍사의 퀵캠 스피어 AF를 뮤직보드를 수직으로 투영하기 위해서 스탠드를 사용하여 고정한다. 웹캠에서 수신 받은 영상을 처리하고 미디 인터페이스를 이용하여 미디 트랙을 출력한 PC는 Intel Quad Core 2.66Ghz CPU와 4기가의 램을 탑재하였다. 선택된 미디 신호를 출력하여 전자악기를 연주하기 위한 미디 기기로는 코르그사의 트라이톤(76건반) 신디사이저를 사용하였으며, 미디 신디사이저와 PC와의 연결을 위한 미디 인터페이스로는 이고 시스템의 M4U 제품을 사용하였다.



그림 10. 제안 시스템의 하드웨어 구성  
Fig 10. Hardware Configuration of the Proposed System

4.1.2 제안 시스템 구현

제안 시스템의 마커 인식부는 FLARToolkit 2.5.2기반으로 하고, Adobe사의 Flash Builder 4를 이용하여 Action Script를 기반으로 개발한다[11][12]. 캠을 통해서 획득된 영상을 기반으로 인식된 정보는 3.3에서 제안한 (C)와 같은 형태로 생성되고 이를 내부 소켓을 이용하여 Microsoft .Net 3.5 기반의 C#으로 개발된 규칙 검색부 데몬으로 전달된다. 규칙 검색부는 3.3에서 제시한 (E)와 같은 형태의 매핑룰 정보에서 해당 뮤직 마커 기반의 정보를 3.3에서 제시한 (D)와 같은 형태로 추론하고 이를 변환한다.

미디어 재생부는 획득된 정보를 가지고 테스트를 위해 미리 작성된 미디 파일에서 해당 트랙 정보를 변경하는데, 이를 위해서 C# MIDI Toolkit을 기반으로 해당 채널의 미디 메시지를 변경하고 파일을 재생하여 음악을 출력한다[13].

4.2 제안 시스템 성능 테스트

제안 시스템의 성능을 평가하기 위하여, 캠의 해상도와 초당 처리 프레임 수와 뮤직 마커의 개수에 따라서 이를 처리하는데 걸린 시스템 내부의 소요 시간을 측정한다.

단계별 소요 시간 측정을 위해서 [그림 11]에서 제시된 시스템 흐름 과정에서 뮤직 마커의 영상을 획득한 후, 미디 파일을 변경하여 출력하기 전까지 소요된 시간을 측정하는 방식으로 각 단계별로 로그를 파일에 저장하고 이와 더불어 단계별 시각정보를 저장하여 단계별 소요 시간을 측정한다. 마커의 위치 파악 단계에서부터 해당 미디 파일을 출력하기 전까지의 시간 차이를 시스템의 지터(Jitter)로 설정한다.



그림 11. 시스템 흐름도  
Fig 11. System FlowChart

이와 더불어 캠의 해상도에 따라서 발생하는 뮤직 마커의 미인식이나 오인식 비율을 측정한다. 뮤직 마커의 미인식은 캠을 통해서 획득된 영상의 마커를 인식하지 못하는 경우를 의미하며, 오인식은 마커를 다른 정보로 잘못 인식하는 경우를 의미한다.

표 4. 캠 설정에 따른 Jitter와 인식률

Table 4. Jitter and success ratio according to CAM configuration

칼라	해상도	초당 처리	Jitter(ms) / 마커수				인식률
			1	2	3	4	
YUV	320 X 240	1	30	45	70	90	70%
		2	100	120	160	190	65%
		3	160	190	230	270	55%
	640 X 480	1	80	110	145	180	90%
		2	200	240	310	390	77%
		3	290	340	390	450	70%
	800 X 600	1	120	160	240	340	100%
		2	280	330	390	450	100%
		3	420	500	560	710	100%
	1280 X 760	1	900	1100	1450	1600	100%
		2	1100	1270	-	-	100%
		3	-	-	-	-	100%
1600 X 1200	1	3500	-	-	-	100%	
	2	-	-	-	-	-	
	3	-	-	-	-	-	

[표 4]는 다양한 해상도와 초당 처리 프레임 설정 값에 따른 지터와 인식률에 대한 결과 값을 의미한다. 지터는 ms 값을 기준으로 측정하였으며 인식률은 모두 인식된 경우를 100%로 산출하였을 때, 발생된 오인식과 미인식 비율의 합을 의미한다. 결과와 같이 해상도가 1280×760이 넘어가는 경우에는 마커의 움직임에 따른 응답속도가 수십 초를 넘어가기 때문에 측정하기 어려웠으며 800×600의 해상도일 때 가장 높은 인식률과 사람이 차이를 분간할 수 없을 만큼의 지터가 발생하였다. 뮤직 마커의 개수가 많아지게 그만큼 소요 시간이 오래 걸리기 때문에 초당 2장 정도의 영상을 처리한 경우에 500ms내에서 시스템 운영이 가능하였다.

4.3 관련 연구와의 비교 분석

관련 연구와의 장단점을 구분하기 위해서 [표 5]와 같이 4 가지 항목에 대해서 비교분석하였다. 평가 방법은 20명의 사용자를 대상으로 각 시스템이 작동하는 개념을 설명하고 시연 화면을 동영상으로 보여주어, 각 시스템별로 상대 비교 점수를 상/중/하로 구분하여 평가하였다. 평가에 항목으로 첫째, 직관성은 시스템에서 사용되는 인터페이스가 사용자 측면에서 얼마나 이해하기 쉬운지를 측정한다. 둘째, 정확성은 마커의 인식결과에 따라서 이를 명확하게 구분할 수 있는지를 평가한다. 셋째, 용이성은 사용자가 체감할 때, 사용자가 특별

한 악기나 음악에 대한 교육없이 시스템을 사용할 수 있는지에 대한 평가 항목이다. 넷째, 현실성은 제안 시스템이 상대적으로 얼마나 현실감있는 증감 효과를 사용자에게 주는지 평가하는 항목이다.

표 5.항목별 상대 평가 결과  
Table 5. Relative Evaluation

관련 연구	직관성	정확성	용이성	현실성
Poupyrev 연구	상	상	중	상
Costanza 연구	중	상	중	상
KIA 연구	상	하	중	중
Liarokapis 연구	중	중	상	중
Motokawa 연구	중	중	상	상
Smith 연구	상	상	상	상
제안 연구	상	상	상	상

평가 항목에서 상대적으로 높은 점수를 획득한 것은 직관적으로 기록된 마커를 이용하여 마커의 종류와 배치만으로도 공감각적인 음악의 편집 및 재생 효과를 얻을 수 있기 때문인 것으로 판단된다.

### V. 결론 및 향후 연구

본 논문에서는 직관적인 뮤직 마커를 기반으로 특별한 음악에 대한 지식 없이도 뮤직 마커를 이용하여 악기를 배치하고, 마커에 해당하는 악기와 반복구간을 설정함으로써 음악을 편집하고 재생할 수 있는 증강 현실 기반의 실시간 음악 편집 및 재생 시스템을 제안하였다. 제안 시스템은 뮤직 보드를 기반으로 뮤직 마커의 위치에 따른 방향과 볼륨에 대한 설정 값을 이용함으로써 음악 재생 과정에서 공간 효과를 적용하였다. 이와 더불어 뮤직 마커를 인식하고 획득한 정보를 메타데이터 형태로 정형화하고 정방향 추론을 통해서 미리 정의된 MIDI 파일의 트랙과 매핑할 수 있는 인프라를 제공하여 다른 MIDI 파일이나 매핑 정보를 변경해야하는 경우, 시스템을 재개발하지 않고 메타데이터의 수정만으로도 확장이 가능하다.

향후에는 제안된 뮤직 마커를 기반으로 해당 뮤직 마커에 해당하는 악기나 악기를 연주하는 연주자의 모습을 3D 그래픽 영상을 오버레이하여 청각적인 효과와 더불어 시각적인 효과를 적용할 수 있는 시스템에 대한 연구가 필요하다.

### 참고문헌

- [1] C. Roads, "Computer Music Tutorial", MIT Press, 1995.
- [2] Poupyrev, I., R. Berry, M. Billinghurst, H. Kato, K. Nakao, L. Baldwin, and J. Kurumisawa, "Augmented Reality Interface for Electronic Music Performance," Proc of HCI, pp.805-808, 2001.
- [3] E. Costanza, S. B. Shelley and J. Robinson, "Introducing Audio D-touch : A Tangible User Interface for Music Composition and Performance," Proc of the 6th Conference on Digital Audio Effects, pp.1-5, 2003.
- [4] K. C. NG, "Music via Motion Transdomain Mapping of Motion and Sound for Interactive Performances," Proc of IEEE, Vol. 92, No. 4, pp.645-655, March 2004.
- [5] F. Liarokapis, "Augmented Reality Scenarios for Guitar Learning," in Proc of the International Conference on EG UK Theory and Practice of Computer Graphics, pp.163 - 170, 2005.
- [6] Y. Motokawa and H. Saito, "Support System for Guitar Playing using Augmented Reality Display," ISMIR 2006, pp.243-244, October 2006.
- [7] S. Smith, D. Suh, "A Virtual Harp with Physical String Vibrations in an Augmented Reality Environment," Proc of ASME, Vol 2, pp.1123-1130, September 2007.
- [8] wikipedia, [http://en.wikipedia.org/wiki/General\\_MIDI](http://en.wikipedia.org/wiki/General_MIDI)
- [9] E. Y. Kim, D.Y. Oh "Design and Implementation of Context-Aware Middleware for Distributed Ubiquitous Environments", Journal of the Korea Society of Computer and Information, Vol. 11, No. 5, pp.105-114, November, 2006.
- [10] J. K. Lee, "Principle and Development of Expert

System”, BubYoung Publishing Company, pp.62-75, August, 1998.

[11] SPARK Project, <http://www.libspark.org/wiki/saqoo-sha/FLARToolKit>

[12] Adobe Flex, <http://www.adobe.com/kr/products/flex/>

[13] C# MIDI Toolkit, <http://www.codeproject.com/KB/audio-video/MIDIToolkit.aspx>

## 저 자 소 개



### 김 은 영

1987 : 숙명여자대학교 전산학과  
이학사.

1991 : 숙명여자대학교 전산학과  
이학석사.

2001 : 숭실대학교 컴퓨터학과 공학  
박사

현 재 : 안산공과대학 멀티미디어컨  
텐츠과 교수

관심분야 : 인터넷 표준, 멀티미디어  
컨텐츠

Email : kispia@empal.com



### 오 동 열

1999 : 경희대학교 전자계산학과  
공학사.

2001 : 숭실대학교 컴퓨터학과  
공학석사.

2010 : 숭실대학교 컴퓨터공학과  
공학박사

현 재 : 인젠트 연구개발부 책임연구원

관심분야 : 증강현실, 영상처리

Email : javarian@gmail.com