

멀티미디어 콘텐츠를 위한  
청각데이터의 시각화 디자인에 관한 연구

A Study on Auditory Data Visualization Design for  
Multimedia Contents

주저자: 홍성대 (Hong Sung-dae)

중앙대학교 영상공학과 예술공학전공

공동저자: 박진완 (Park Jin-wan)

중앙대학교 영상공학과 예술공학

서 론

- 1-1 연구의 배경 및 필요성
- 1-2 연구의 목적 및 필요성

2. 멀티미디어 콘텐츠와 커뮤니케이션 흐름의 변화

- 2.1 멀티미디어 콘텐츠의 개념과 특성
- 2.2 멀티미디어 콘텐츠에서의 디자인
- 2.3 커뮤니케이션 패러다임의 전환
- 2.4 오디오 관련 작품사례 분석

3. 인터랙티브 디자인에 필요한 오디오 콘텐츠의 분석

- 3.1 음향신호의 분석
- 3.2 음향값의 시각화
- 3.3 주파수
- 3.4 프로그램 구현
- 3.5 실시간 구현 방법

4. 인터랙티브 디자인 및 고찰

- 4.1 플래시를 이용한 제작구현
- 4.2 디렉터를 이용한 제작구현
- 4.3 디렉터를 이용한 리얼타임 제작구현

5. 결 론

참고문헌

(要約)

디지털 기술의 발달로 점차 디자인(예술), 미디어, 과학 분야는 개인화와 함께 맞춤형(customization) 추세로 가고 있다. 기존 대중매체는 기술 및 경제적 제한요인으로 인하여 대중에게 일방적으로 뿌려지는 형태를 띠었고, 예술작품 또한 일정한 전시 공간 또는 무대에서 일방적으로 관객에게 전달하는 방식으로 소통 해왔다. 그러나 현재의 상품(또는 예술품)은 관객, 시청자 또는 소비자의 직접적인 참여가 가능하다는 점이 주목할 특징이다. 각각 미디어나 예술을 소비하는 개인의 취향에 따라 상품을 다르게 만들 수 있게 되었는데 기술의 핵심은 '인터랙티브 기술'이다.

이에 본 연구는 멀티미디어 콘텐츠 개념을 '정보혁명' 측면에서 살펴보고, 다음으로 문제해결을 위한 분석의 과정으로 시청각 인터랙티브 디자인의 접근을 상호작용의 정보 구조로 모색해 보며, 문제 해결의 방안으로서 오디오에서 시각적으로 표현 될 수 있는 음량 및 주파수를 추출하는 프로세스를 제시하여 오디오의 추출값을 비주얼로 표현하는 연구 과제로 최종 결과물은 디자이너의 입장에서 작품 형태로 제시하였다. 이와 같은 연구를 통하여 디지털 콘텐츠에서의 인터랙티브 디자인이 우리시대에 더욱 긴밀하고 새로운 소통의 방법을 제시하며, 더 넓은 의미의 "상호소통"을 촉진시킬 수 있다면, 디자인이 갖는 본래의 기능들에 일조할 뿐 만 아니라, 미래 디자인의 예측에 있어서도 더욱 다양한 방법과 비전을 제시할 수 있을 것이라고 본다.

(Abstract)

Due to the of evolution of digital technology, trends are moving toward personalization and customization in design (art), media, science.

Existing mass media has been broadcasting to the general public due to technical and economic limitation and art works also communicate one-sidedly with spectators in the gallery or stage. But nowadays, it is possible for spectators to participate directly. We can make different products depending on the tastes of individuals who demand media or art. The essence of technology which makes it possible is 'interactive technology'.

A goal of this research is to find out the true nature of the interactive design in multimedia contents and find the course of interactive communication design research. In this paper, we pass through two stages to solve this kind of problem. At first, we studied the concept of multimedia contents from the aspect of information revolution. Next, we decided our research topic to be 'visual reacting with audio' and made audio-visual art work as graphic designers. Through this research we can find the possibility to promote 'communication' in a broad sense, with appropriate interactive design.

(Keyword)

Interactive Design, Amplitude, Real Time

# 1. 서론

## 1-1. 연구의 배경 및 필요성

디지털 혁명은 과거 산업 혁명으로 발생한 산업과 문화의 발전 이상으로, 엄청난 변화를 야기할 것으로 예상된다. 그러나 산업혁명이 특정 국가를 중심으로 비교적 서서히 삶의 양식에 영향을 미친 것에 반하여, 디지털 혁명은 초고속 정보망과 첨단 컴퓨터 기술을 통하여 동시에 확산되는 엄청난 영향력을 지니고 있다는 점에서 그 혁명의 정도가 훨씬 크다. 또한 컴퓨터를 이용한 첨단 정보 기술은 19세기 사진 기술이 표현예술계에 미친 영향보다 큰 충격을 예술과 문화계에 주고 있다.

디지털 기술의 발달로 점차 디자인(예술), 미디어, 과학 분야에서는 개인화와 함께 맞춤형(customization) 추세로 가고 있다.

기존 대중매체는 기술 및 경제적 제한요인으로 인하여 다수 대중에게 일방적으로 뿌려지는 형태를 띠었고, 예술작품 또한 일정한 전시 공간 또는 무대에서 일방적으로 관객에게 전달하는 방식으로 소통 해왔던데 반해 현대 디지털 매체는 기존 제한을 뛰어넘어 다름 아닌 관객, 시청자 또는 소비자의 직접참여가 가능해졌다. 각각 미디어나 예술을 소비하는 개인의 취향에 따라 상품(또는 예술품)을 다르게 만들 수 있게 되었는데 그 기술의 핵심은 '인터랙티브 기술'이다.

본 연구에서는 인터랙티브 디자인의 사례를 살펴보고 다중매체 속에서 소리를 시각화하기 위한 음량 값을 다양하게 추출하여 소리에 반응하는 시각적 표현과 효과를 멀티미디어 저작 소프트웨어로써 형태를 제안하며, 보다 더 효율적으로 제작하기 위하여 멀티미디어 콘텐츠의 새로운 Interactivity를 모색한다.

## 1-2. 연구의 범위와 방법

본 연구에서는 멀티미디어 콘텐츠의 한 분야인 인터랙티브 디자인을 개발하기 위하여 소리에 연동되는 시각적 표현방법을 연구하였다. 이를 위해서는 소리에 대한 전반적인 이해가 필요하다, 따라서 구현과 관련된 프로세스 진행은 Macromedia사의 소프트웨어에 바탕을 둔 것임을 밝힌다.

본인이 주장하는 인터랙티브 디자인의 작품을 제작하기 위해서는 두 가지의 중요한 매체의 연구가 필요하다. 우선 소리라는 매체를 통하여 시각화할 수 있는 값을 추출하는 것이며, 또 다른 한 가지는 멀티미디어 개발 소프트웨어인 Macromedia 사의 Director(lingo script) 및 Flash(Action Script)를 통한 시각적 표현매체에 대한 연구이다. 이러한 두 가지 이상의 매체가 혼합되어 개발되어진 방법은 본인의 인터랙티브 디자인의 작품형태로 구성된다.

# 2. 멀티미디어 콘텐츠와 커뮤니케이션 흐름의 변화

## 2-1. 멀티미디어 콘텐츠의 개념과 특성

멀티미디어 PC, 멀티미디어 음악, 멀티미디어 잡지, 멀티미디어 교실, 멀티미디어 박물관, 모든 명사의 앞에만 붙이면 그럴듯해 보이는 '멀티미디어' 라는 단어는 어디선가 갑자기 나타나 이제는 컴퓨터라는 기계와 함께 생활의 일부분이 되어버렸다.

멀티미디어라는 단어는 '여러 가지'(multi)와 '매체'(media)라는 두 단어의 합성어로 두 단어의 의미에서 추정하면 여러 가지 매체, 혹은 다양한 매체의 사용으로 볼 수 있다. 그러나 '미디어' 라는 단어가 표현할 수 있는 의미의 스펙트럼은 매우 넓으므로 그 구체적 의미에 대해 생각해 볼 필요가 있다. 미디어는 매스미디어라는 단어에서 볼 수 있듯이 신문, 방송, 잡지 등 전달자라는 의미를 가진 무형의 매체를 의미하는 경우가 있으며 카세트 테이프나 비디오 테이프를 제작하는 회사명에서 자주 볼 수 있듯이 내용물을 담을 수 있는 물리적 매체를 의미하기도 한다. 그러나 멀티미디어라는 단어에 포함된 미디어의 의미는 정보를 표현하는 방법을 말한다. 엘빈 토플러는 현대 사회는 권력의 중심이 자본에서 지식으로 이동하는 사회라고 했다. 지식이 권력인 사회, 즉 '아는 것이 힘'인 정보혁명의 사회인 것이다.

오늘날의 정보혁명이 있기 이전에도 정보는 존재해 왔다. 신문, 방송, 잡지 등의 매스미디어(mass media)방식을 통해 수많은 데이터가 이동해 왔고, 인간은 이러한 데이터 중에서 자신의 필요에 적합한 것을 수집하여 정보로 활용할 수 있었다. 그러나 오늘날의 정보혁명에는 이러한 과거의 정보이용과는 다른 특징이 있다.

정보전달의 효율성을 높이기 위한 경험적 인터페이스 즉 인터랙티브의 발전이다. 이러한 정보혁명의 특징은 인간이 정보를 활용함에 있어 일방적인 강요보다는 상호작용에 의해 정보의 취사선택을 선호하게 하였고, 그 결과 정보제공 방법의 변화를 초래하게 하였다. 산업사회에서의 인간은 텔레비전을 시청하거나 신문을 읽는 것과 같이 매스미디어에 의해 제공된 정보를 일방적으로 수용하는 입장이었다. 하지만 정보혁명은 정보를 한곳에 모아놓고 인간으로 하여금 필요한 정보를 선택하도록 하였다. 이러한 과정에서 정보를 표현하는 사용자와의 인터페이스 문제가 중요하게 되었고 특히 제공 방법에 대한 네비게이션(navigation)의 문제 즉 사용자가 그 정보를 어떠한 방식에 의해 획득할 수 있도록 할 것인가 하는 문제가 매우 중요하게 대두되었다. 즉 정보가 하나의 제품으로 인식되면서, 정보를 보다 많이 그리고 보다 효율적으로 사용할 수 있도록 하기 위하여 사용자의 경험에 친숙한 제공방법을 결정하는 문제가 매우 중요하게 된 것이다.

기존의 콘텐츠	멀티미디어 콘텐츠
책,신문,라디오,영화,등의 old media로 제공되는 콘텐츠	총체적인 매체를 활용하여 재창출시키는 제반 내용물
아날로그	디지털
단방향으로 제공	쌍방향으로 제공
콘텐츠의 생산자와 수용자가 명확히 구분됨	정보이용자가 동시에 정보제공
시,공간의 제약을 받음	시,공간의 제약 없이 제공
Up Date의 비용이 크다	Up Date의 비용이 적고 쉽다.
순차적인 방법에 의해 정보 제공	정보의 습득과정이 비순차적

[표 1] 기존의 콘텐츠와 멀티미디어 콘텐츠 비교

정보혁명이라는 것은 사회를 움직이는 원동력이 농업이나 자

1) Alvin Toffler : 1928년~ 뉴욕 출생. 저널리스트. 미래학자. 「미래(未來)」지의 부편집자로 활동. 뉴욕의 뉴스쿨 교수. 코넬대학 초빙교수 역임

본이 아닌 정보로부터 시작된다는 것을 의미한다. 이 정보를 이루고 있는 것, 혹은 정보 그 자체가 바로 콘텐츠라고 할 수 있으므로 이렇게 콘텐츠가 폭증을 하는 것은 오히려 자연스러운 현상이다. 그렇다고 콘텐츠가 갑자기 솟아난 것은 아니다. 이미 인류 역사와 함께 수천 년을 함께 해왔던 출판, 각종 교육, 최근의 라디오, TV, 방송, 영화 등 인간이 창조해낸 모든 지식과 상상, 경험이 바로 콘텐츠이지만 정보혁명을 맞이하면서 우리가 거론하는 것은 가공되고, 저장되고, 쉽게 전송될 수 있는 콘텐츠, 즉 디지털로서의 콘텐츠에 중점을 두고 있다.<sup>2)</sup> 콘텐츠라는 말을 새삼스럽게 만들어 낼 필요가 생긴 것은 PC가 표현의 도구로 보급되기 시작했기 때문이다. 워드프로세스 소프트웨어, 동작 재생 소프트웨어 등 PC 소프트웨어와 달리 이들 소프트웨어를 이용하여 창조되어진 문장, 그림, 영상 등을 말한다. 그러므로 콘텐츠 자체는 결코 새롭게 특별한 것은 아니다. 그러나 그것이 멀티미디어 콘텐츠로 되면 특별한 것이 된다. PC를 표현의 도구로 사용하게 되었고 게다가 저렴하게 구입할 수 있게 되었기 때문에, 이전에는 정보를 일방적으로 받아들이는 쪽이었던 일반 개인이 이제는 정보를 보낼 수도 있게 된 것이다.

## 2-2. 멀티미디어 콘텐츠에서의 디자인

콘텐츠는 구체적 대상(object)을 생산하는 것이 아니라 내용 또는 방법을 만들어 내는 것이기 때문에, 디자인 콘셉을 확립하는 단계에서 '어떠한 특성을 가진 작품을 만들 것'이라는 기준을 세우는 것만으로는 바람직한 디자인 결과를 얻기가 어렵다는 점이다. '어떠한(what) 정보를 만들어 내는가' 하는 문제보다는 '어떻게(how) 효과적으로 검색하고 표현할 수 있는가'라는 목표를 가지고 그 방법을 추상적으로 설정하는 문제가 더욱 중요하기 때문이다. 따라서 정보디자인의 컨셉 설정 단계에서는 과거와는 다른 새로운 디자인 접근 방안이 모색되어야 한다.

첫째, 논리적 특성에서의 접근 방법이다. 전통적 디자인에서는 제품의 사용 방법과 관련된 기능적인 문제를 물리적 측면에서 고려하는 수준으로 디자인에 접근하지만, 콘텐츠에서는 제품의 쌍방향 대화기능(interactivity)이 제품의 성패를 결정짓는 가장 중요한 요소로 작용하기 때문에 제품 개발의 전 과정에서 이것이 지속적으로 검토될 수 있는 방안이 마련되어야 한다는 점이다. 즉 단지 물리적 기능성을 일차적이고 고려하는 지금까지의 디자인과는 달리, 경험적, 방법적 기능성이 디자인 작업의 전 과정을 통해 중요하게 다루어져야 한다는 점을 명확히 해야 하는 것이다.

둘째, 표현 및 인지적 특성에서 접근 방법이다. 디자인이 과거에는 시각적 측면에서의 조형성 창조의 역할에 충실하였다면 콘텐츠에서는, 비록 그 비중을 산술적으로 나누는 것이 타당한 것은 아닐지라도, 이러한 시각적 측면의 고려가 다양한 특징들 중의 한 부분으로 다루어져야 한다는 점이다. 즉 산업 제품이 시각적 특성이 강조를 통해 제품의 판매 경쟁력을 높였다면, 콘텐츠는 시각적인 측면과 더불어 청각적인 측면을 함께 고려하여 제품의 사용 경쟁력을 강화하고 이를 통한 기

능성의 강조와 즐거움(entertainment)을 부여할 수 있는 디자인이 진행되어야 하는 것이다. 이를 위하여 앞으로의 디자인에서는 시각적 요소인 영상과 그래픽에 관련된 내용뿐만 아니라 문안의 내용이나 음향, 나레이션과 같은 비시각적 요소에 대한 디자인 과정도 포함되어야 한다.

셋째 생산적 특성에 따른 디자인의 접근 방법이다. 전통적인 디자인은 디자인과 기타 부문이 분리되어 디자인 내부의 문제에 집착해 왔지만 정보디자인 프로세스는 설계의 과정이 디자인의 중요 역할로 존재하게 되고 생산에도 디자인의 프로세스를 새로이 마련해야 할 필요성이 나타나게 될 것이다. 즉 디자인이 생산의 한 부분을 담당하는 것이 아니라 생산의 전체를 주도할 수 있는 역량을 가져야 하는 것이다.<sup>3)</sup>

## 2-3. 커뮤니케이션 패러다임의 전환

이제까지 신문, 방송, 라디오, 영화, 출판 등으로 구분되어 있는 개별 미디어들이 하나의 플랫폼으로 통합되어가고 있는 추세를 볼 수 있다. 여기에 덧붙여 홈쇼핑, 홈뱅킹, 원격진료, 화상회의 등의 서비스산업이 인터넷으로 통합되어 그야말로 현실사회의 다양한 기능들이 멀티미디어 콘텐츠에 편입되고 있다. 미국을 중심으로 건설되고 있는 초고속 정보통신 기반 구축사업의 어플리케이션(application) 영역에서 이러한 서비스들은 주요사업으로 이미 일정수준에 올라 있으며 향후 이러한 통합 또는 혼합모델(hybrid model)의 추세는 더욱 가속 될 것으로 예측할 수 있다. 이러한 예측이 가능한 것은 CD-I나 VOD, 양방향케이블TV등 차세대의 미디어 서비스 영역들이 인터넷을 포함하는 서비스로 그 영역을 확장하거나 인터넷상에 직접 연결되는 하나의 인터넷 서비스로 전환되고 있다는 사실에서 볼 수 있다. 현재까지의 멀티미디어 시스템은 컴퓨터와 방송, 통신망 측면에서 각각 계속적인 발전을 하고 있지만, 관련 기업간의 흡수와 합병, 기술 제휴 등으로 인해 결국에는 이 세 분야가 통합도니 다양한 형태의 서비스를 제공하는 멀티미디어 시스템으로 개발될 것이다. 이상 살펴본 것과 같이 멀티미디어 콘텐츠가 커뮤니케이션의 패러다임을 전환하는 새로운 계기가 되고 있는 가운데, 인터넷이 인쇄매체, 라디오, TV 다음의 제 4의 미디어인가, 아니면 다른 모든 미디어를 통합하는 토털 미디어 즉 미디어의 통합체로 발전해 나갈 것인가 하는 문제는 의미심장하다. 시각디자인은 기본적으로 미디어의 디자인에서 출발한 것이며 앞으로 그 그러할 것이다. 그러나 이와 같은 커뮤니케이션 네트워크를 통한 미디어의 통합추세와 관련하여 시각디자인의 정체성에 어떤 변화가 올 것인가를 생각해 볼 필요가 있다. 왜냐하면 멀티미디어에서 커뮤니케이션의 질을 높이는 일은 매우 중요하기 때문에 누구나 해야 할 일이며 따라서 시각적인 정에 대한 문제는 더 이상 전문가의 배타적인 영역이라고 할 수 없게 될 것이기 때문이다. 모든 조형 활동은 문제의 해결안 또는 인간과 세계에 대한 해석, 이미지네이션, 비전 등의 정신세계 현실 속에 구현하는, 즉 상상의 세계의 인터페이스로서 현실공간에서 이루어져왔다. 그러나 급격하게 변화하는 멀티미디어 세계커뮤니케이션

2) 박현제, 콘텐츠 산업과 디자인 벤처, 월간디자인.1999.1, p.104

3) 홍원기, Strategic Interactive Communication Design Research for the Multimedia Contents, 경원대학교, 1999, p.16

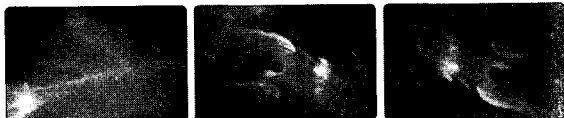
패러다임을 빠른 속도로 전환시키고 있다. 기존의 미디어에 익숙해있던 커뮤니케이션 디자이너에게는 엄청난 도전이 아닐 수 없다.

## 2.4 오디오 관련 작품사례 분석

인터랙티브 작품사례연구는 소리에 관련된 작품을 시대적으로 연구하였으며, 또한 현대에 들어서 MIT 미디어랩에서 연구하는 작품과 본인이 연구하는 소리에 반응하는 작품으로 토대로 사례 연구를 하였다.

### · 1911 Alexander Scriabin(1872-1915)

러시아 작곡가Alexander Scriabin는 그의 교향곡 Poem of Fire (Prometheus)에서 빛의 색상들을 변화시키는 부속 장치들을 고안하였다. 그가 Tastiera per Luce라고 이름 붙인 장치는 키보드 기반 칼라 오르간이라고 불리었고 아마도 Rimington의 악기를 참고한 것으로 보인다. 그는 관객들이 흰색 옷을 입기를 바랐는데 그 이유는 관객의 몸만이 아니라 전체 방에 색상들이 프로젝션 될 것이기 때문이다.



[그림 1] Tastiera per Luce

### · MIT Media Lab 작품

시청각 환경 슈트(AVES)는 사람에게 실시간에 만들고, 추상적인 애니메이션과 합성 소리를 수행하는 시스템은 5대의 상호 교환적인 시스템이다. 각 환경은 배우기 쉽다. 또한 시각적으로 흥미 있고 무한히 변화하기 쉽고 개인적으로 표정이 풍부한 실행을 낳을 수 있는 인터페이스를 디자인한 실험적 시도이다.



[그림 2] Scribble Concert. Golan Levin, MIT Media Laboratory, 2000

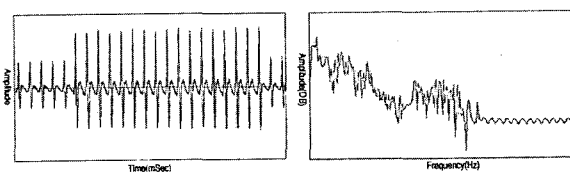
## 3. 시각화에 필요한 오디오 콘텐츠의 분석

### 3-1. 음향신호의 분석

음향 신호의 분석은 크게 ①시간 영역에서 표현하는 방법과 ②주파수 영역에서 표현하는 방법으로 나눌 수 있다. 시간 영역에서의 표현은 그림 3의 a)와 같이 가로축에 시간이, 세로

축에 진폭(크기)이 나타나는 그래프로 표시된다. 즉, 오실로스코프로 파형을 관측하는 것이 일반적인 시간 영역의 표현방법이 된다. 주파수 영역에서의 표현은 [그림 3]의 b)와 같이 가로축에 주파수가, 세로축에 진폭(크기)이 나타나는 그래프로 표시된다. 즉, 스펙트럼 분석기로 관측하는 것이 일반적인 주파수 영역의 표현방법이 된다. 디지털 사운드 데이터는 오실로스코프나 스펙트럼 분석기와 같은 장비 없이도, 일반적인 PC상에서 돌아가는 Cool Edit, Sound Forge같은 음향 편집 프로그램에서도 정확한 측정이 가능하다.

어떠한 신호라도 시간 영역과 주파수 영역으로 분석하여 나타낼 수 있다. 이렇게 분석된 신호는 [그림 3]과 같이 시각적으로 표현 될 수 있으며, 또한 분석된 데이터를 기반으로 인터랙티브 디자인에 적합하게 변형하여 응용할 수 있다.



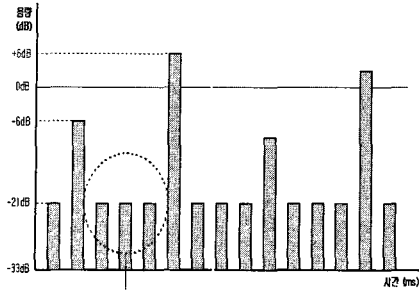
(a)시간 파형 (b)스펙트럼  
[그림 3] 음향 데이터의 시간 파형과 스펙트럼

### 3-2. 음향값의 시각화 방법

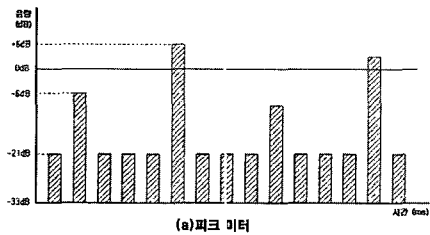
소리를 시각화 하는데 있어서 가장 직관적으로 사용할 수 있는 것은 음량이다. 음량이란 소리의 크기를 뜻하며, 물리적인 측면에서는 음압의 세기, 즉 압력의 크기로 이해 할 수 있다. 그리고 디지털 사운드 포맷이라면 음량은 해당 샘플의 크기로 이해할 수 있다. 음량은 그림 1에서와 같이 시간에 따른 소리의 크기로 표현 할 수 있다. 하지만 인간의 시각과 청각에 적용시키기 위해서는 [그림 3]의 데이터를 그대로 사용할 수는 없다. 가장 일반적으로 사용되는 CD나 WAV 파일의 샘플링 레이트는 44100Hz이다. 즉, 초당 44100개의 음량 데이터를 뽑아 올 수 있으나, 인간의 시각과 청각이 그에 맞게 적용하지 못하기 때문이다. 인간의 시각에 적용하기 위해서는 적게는 초당 15개, 많게는 60개 정도의 대표되는 데이터들로 줄여서 표현해야 한다.

초당 44100개의 데이터를 초당 30개의 데이터로 줄이기 위해서는 물리적인 소리의 크기 변화에 따라 인간의 청각 특성이 어떻게 변화되는지에 대해서 알아야 한다. [그림 4]에서의 한 막대는 50ms 정도의 지속시간을 가진다. [그림 5]의 a)는 음량 측정용 장비인 피크미터를 사용했을 때의 음량 변화를 나타낸 것이고, b)는 인간이 들었을 때의 음량 변화를 나타낸 것이다. 그림에서 알 수 있듯이, 측정 장비는 측정할 수 있는 순간적인 음량의 변화를 인간의 귀는 감지하지 못한다. 이는 50msec 정도의 짧은 지속시간을 가진 음량의 변화는 인간의 귀가 감지하지 못하기 때문이다. 일반적으로 인간이 음량의 변화를 인식하기 위해서는 200~300msec 정도의 지속시간이 필요하며, 음량의 변화를 느끼는데 필요한 시간을 인티그레이션 타임(Integration Time) 이라고 한다. 음향 데이터의 왜곡을 막기 위해 사용되는 측정 장비인 피크 미터는 1~10msec 정도의 인티그레이션 타임을 가지며, 인간이 느끼는 음량을 측정하는

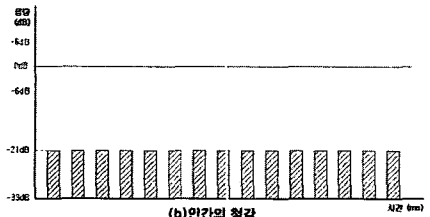
측정 장비인 VU Meter는 300msec 정도의 인티그레이션 타임을 가진다.



[그림 4] 음량 레벨 변화



(a) 피크 미터

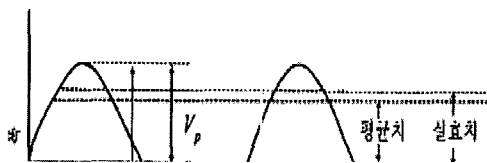


(b) 인간의 청감

[그림 5] 피크미터와 인간의 청감

인티그레이션 타임을 고려한다면, 초당 뽑아올 수 있는 데이터는 3~5개 정도가 된다. 하지만 이 또한 인간의 시각적인 특성에 적합하지 못하여 너무 느린 반응성을 가지게 된다. 즉, 시각적인 측면을 고려한다면 초당 15~30(인티그레이션 타임: 33.3 ~66.7 msec)개 정도의 데이터를 뽑는 것이 적당하다. 이는 인간의 청감 특성을 그대로 반영하고 있지는 않지만, 일반적인 사운드 데이터에서 비슷한 수준의 음량이 200~300msec 정도 지속될 가능성이 크며, 보다 부드러운 시각적인 효과를 기대할 수 있다는 점에서 적합하다고 할 수 있다.

교류의 전압을 측정하는 데는 [그림 6]와 같이 최대값(Vp), Peak to Peak(Vp-p), 평균값(Vave), 실효0값(Vrms)의 4가지 방법이 있다. 음향 데이터를 일종의 교류라고 생각한다면, 교류 전압의 진폭을 표시하는데 사용되는 방법들을 적용시킬 수 있다.



[그림 6] 정현파의 레벨 표기

일반적으로 음량을 계산하는 데는 평균값(Average)과 실효0

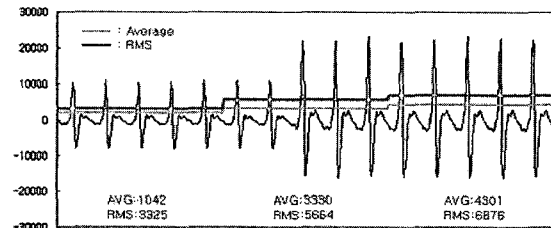
(RMS)을 사용한다. 이렇게 평균값과 RMS를 사용하는 이유는, 일정한 음향이 시간이 지속되어도 계속 같은 레벨로 유지 된다면 인간의 청각은 수학적인 평균값과 동일한 레벨로 느끼지만, 시간에 따라 레벨이 변화하게 되면 인간의 청각이 느끼는 평균 레벨이 수학적인 평균값과는 달라지기 때문이다. 보다 인간이 느끼는 실제적인 평균레벨에 가까운 것이 RMS(Root Mean Square) 방식인 것이다.

여기서 평균값은 식(1)에 의해서, RMS 값은 식(2)의하여 구할 수 있다.

$$AVG = \frac{\sum_{n=1}^N |X_n|}{N} = \frac{(|X_1| + |X_2| + \dots + |X_{N-1}| + |X_N|)}{N} \quad \begin{matrix} N: \text{전체 데이터 개수} \\ X_n: n\text{-에서의 값} \end{matrix} \quad (1)$$

$$RMS = \sqrt{\frac{\sum_{n=1}^N X_n^2}{N}} = \sqrt{\frac{(X_1^2 + X_2^2 + \dots + X_{N-1}^2 + X_N^2)}{N}} \quad \begin{matrix} N: \text{전체 데이터 개수} \\ X_n: n\text{-에서의 값} \end{matrix} \quad (2)$$

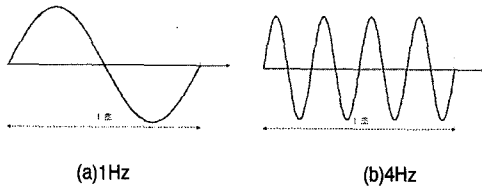
[그림 7]는 실제 사운드 데이터를 가지고 평균값과 RMS 값을 구해 본 것이다. 총 900개의 데이터를 가지고, 300개를 하나의 구간으로 하여 평균값과 RMS를 구하였다. 그림과 같이 평균값과 RMS 값에 차이가 있음을 확인 할 수 있다. 일반적으로 RMS 레벨이 평균값 보다 높은 수치를 보이며 인간의 청각 특성에 더 적합하다고 알려져 있지만, 시각과 청각을 동시에 사용하는 상황에서는 평균치를 사용하는 것이 긍정적인 효과를 가져 올 경우도 있다. 따라서 상황에 맞게 RMS 레벨과 평균 레벨을 적절히 선택하여 사용하여야 한다.



[그림 7] 평균치(Average)와 실효치(RMS)

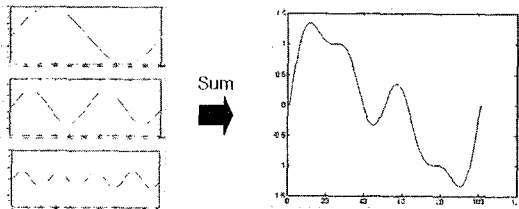
### 3-3. 주파수

주파수란 1초에 몇 번의 주기로 진동하는가를 나타내며, 그 단위는 Hz를 사용한다. 즉, [그림 8] (a)와 같이 1초에 1주기가 완성되면 1Hz 이며, [그림 8] (b)와 같이 1초에 4주기가 완성되면 4Hz 이다. 일반적으로 주파수가 높다는 것은 음이 높다는 것을, 주파수가 낮다는 것은 음이 낮다는 것을 의미하며 어떤 주파수들로 음이 구성되어 있는가에 따라서 느껴지는 음색이 달라진다. 주파수는 소리를 표현하는데 아주 중요한 개념으로, 소리를 시각적으로 표현하는데 있어서도 아주 효과적으로 사용 될 수 있다.

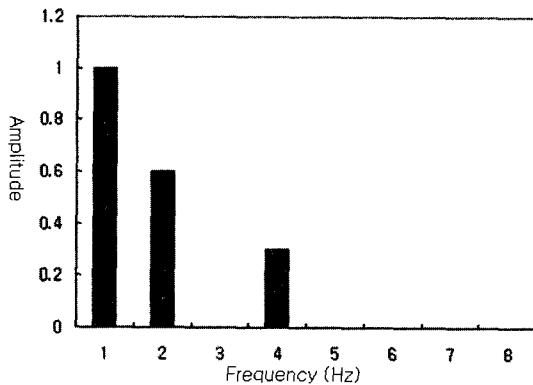


[그림 8] 1Hz와 4Hz

하나의 주파수로만 구성된 음을 순음이라고 하며, 여러 개의 주파수로 구성된 음을 복합음이라 한다. 실제로 순음은 존재하지 않으며, 특정 목적을 위해서 인위적으로 만들어진다. 복합음은 그 음이 어떤 주파수들로 구성되어 있는지, 그리고 각각의 주파수들의 크기는 얼마인지를 분석할 수 있다. [그림 9] (a)는 각각 1Hz(크기 1), 2Hz(크기 0.6), 4Hz(크기 0.3)의 신호를 보여주고 있으며, [그림 9] (b)는 세 개의 신호가 합쳐졌을때의 신호를 보여주고 있다. [그림 9] (b)와 같은 신호는 [그림 10]와 같이 각각의 주파수 성분과 그 크기로 분석 될 수 있다.



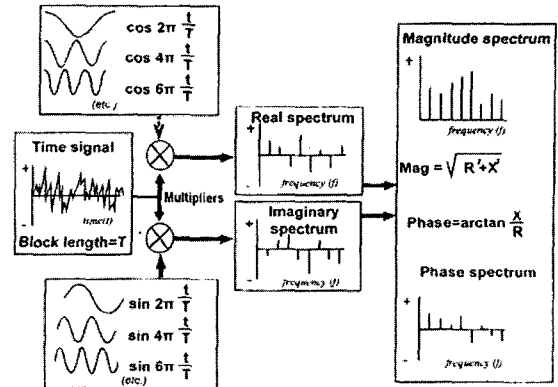
(a) 각각의 신호 (b)합쳐진 신호  
[그림 9] 3개 신호의 합성(1Hz, 2Hz,4Hz)



[그림 10] 주파수 분석

이러한 주파수 분석의 기본은 푸리에 변환(Fourier Transform)이다. 푸리에 변환은 [그림 11]과 같이 시간 영역의 신호에 코사인(cosine)과 사인(sine)을 곱하여서 주파수별로 분해하는 과정이다. 하지만 푸리에 변환을 디지털 음향 신호의 주파수 분석에 그대로 적용시키기는 적절치 않다. 모든 신호는 무한대로 취급할 수 없으며, 대부분의 신호가 주기를 가지고 있으므로 창 함수를 사용하여 일정 부분만을 잘라서 분석하는 방법을 사용한다. 이러한 변환 과정을 디스크리트 푸리에 변환(Discrete Fourier Transform)이라고 하며 그 식은 (3)

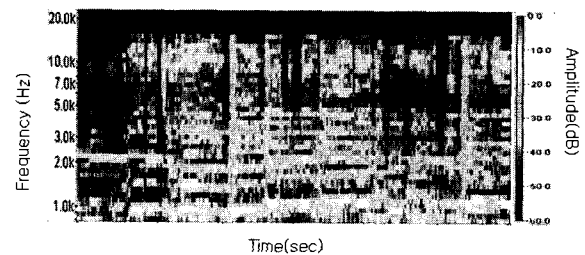
과 같다.



[그림 11] 푸리에 변환

$$X(\omega) = \sum_{n=0}^{L-1} x(n)e^{-j\omega n}, \quad 0 \leq \omega \leq 2\pi, \quad e^{-j\omega n} = \cos \omega n - j \sin \omega n \quad (3)$$

식 (3)을 통해서 DFT(Discrete Fourier Transform)을 하면 실수와 허수 성분을 얻을 수 있으며, 이를 통해서 크기(Magnitude)와 위상(Phase)을 구할 수 있다. 시각적인 표현을 위해서 일반적으로 크기(Magnitude) 성분이 사용된다. 이 때 저주파 쪽에 에너지가 급격하게 몰리게 되므로, 대수관계인 데시벨(dB)을 사용하여 크기를 상대적인 레벨로 나타내는 것이 일반적이다. 최종적으로 음향 데이터를 시간에 따라서 DFT 하게 되면 [그림12]과 같이 시간과 그 시간에서의 주파수별 크기를 얻을 수 있다.



[그림 12] Spectrogram

DFT는 연산량이 많은 변환으로 이를 빠르게 수행하기 위해서 FFT(Fast Fourier Transform)를 사용한다. FFT는 DFT를 빨리 수행하기 위한 알고리즘으로 그 결과는 DFT와 동일하다. 여러 종류의 FFT 알고리즘이 있으며, 공개되어 있는 FFT 소스도 많이 있는 편이다.

DFT를 통해 얻은 데이터를 그대로 시각화에 사용할 수도 있지만, 다시 밴드(Band)를 나누어 사용하는 것이 일반적이다. 1024 포인트 DFT를 수행하면 512개의 데이터를 얻게 되며, 512포인트 DFT를 수행하면 256개의 데이터를 얻게 된다. 즉, DFT 포인트 절반만큼의 데이터를 얻을 수 있다. 이는 시각적인 데이터로 연결하기에는 너무 많은 수이다. 따라서, 적절한 수준의 밴드로 나누어서 사용하게 된다.

밴드를 나누는 방법에도 여러 방법이 있으며, 대표적으로 1) 선형 밴드(Linear Band), 2)로가리즘밴드(Logarithms Band), 3) 옥타브 밴드(Octave Band)가 있다.

1)선형 밴드는 구간을 등간격으로 나누는 것이다. 즉, 20~22050Hz 까지의 주파수 대역을 8 선형 밴드로 나누면 [20~2773], [2774~5527], [5528~8281], [8282~11035], [11036~13789], [13790~16543], [16544~19297], [19298~22050]이 된다.

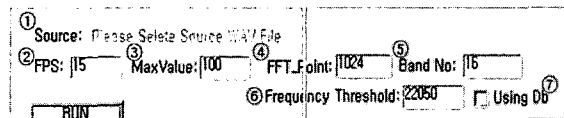
2)로가리즘 밴드는 주파수 대역에 로가리즘을 적용하여, 점점 폭이 넓어지는 밴드를 사용하는 것이다. 이와 같은 로가리즘 밴드를 사용하는 것은 일반적인 음향 신호에서 저주파수 쪽에 에너지가 집중되기 때문이다. 로가리즘 밴드를 사용하면 아래쪽 밴드에 집중되는 에너지를 위쪽의 밴드로도 분산시킬 수 있다는 장점이 있다. 20~22050Hz 까지의 주파수 대역을 8 로가리즘 밴드로 나누면 [20~1213], [1214~3106], [3107~5495], [5496~8268], [8269~11355], [11356~14708], [14709~18291], [18292~22050]이 된다.

3)옥타브 밴드는 옥타브의 개념을 도입한 것이다. 옥타브란 주파수가 1:2가 되는 것을 말하며 음악에서 많이 사용되는 개념이다. 옥타브는 식 (4)로 구할 수 있다.

$$\frac{f_2}{f_1} = 2^n \text{ 수}; f_1: \text{최저한계주파수}, f_2: \text{최고한계 주파수}, n \text{ 옥타브 수} \quad (4)$$

선형 밴드, 로가리즘 밴드, 옥타브 밴드 모두 각각의 장/단점을 가지고 있다. 따라서 소리가 가지고 있는 주파수 특성에 맞게 적절한 밴드를 택하여 사용하여야 한다. 상황에 따라서는 DFT를 통해서 구한 모든 데이터를 그대로 사용하는 것도 좋은 효과를 얻을 수 있다.

### 3-4. 프로그램 구현



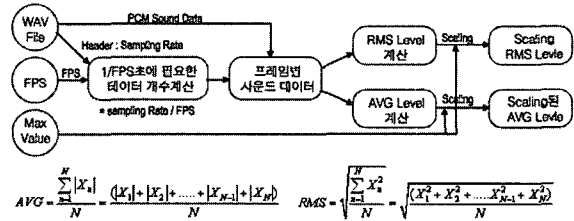
[그림 13] 구현 프로그램 모습

프로그램은 대상 사운드 파일을 입력받고, ②~⑦까지의 파라미터를 입력받아 음량과 주파수를 동시에 분석하도록 구현되어 있다. 사운드 기본 포맷인 WAVE 파일을 입력 받으며, 스테레오로 분석할 수 있다. MFCC 기반으로 구현되어 있다. 기본적으로 ①에서 WAVE 파일을 입력받고, ②에서 FPS(Frame per Sec)와 ③에서 최대값을 입력 받으면 음량의 추출이 가능하다.

우선 입력된 FPS(Frame per Sec), 즉 초당 프레임 수와 대상 웨이브파일의 샘플링 레이트를 기반으로 매 프레임(1/FPS 초) 마다 필요한 사운드 데이터의 개수를 식(5)에 의해서 구한다.

$$F_s = 2^n \text{ 수}; f_s: \text{최저한계주파수}, f_c: \text{최고한계 주파수}, n \text{ 옥타브 수} \quad (5)$$

사운드 데이터를 프레임마다 필요한 개수에 맞게 나누고 식 (1)과 식(2)에 의해서 평균값과 RMS 값을 동시에 계산한다. 마지막으로 계산된 평균값과 RMS 값을 입력된 최대값에 맞게 스케일링(Scaling)하여 최종적으로 각 프레임에 해당하는 평균값과 RMS 값을 출력한다. 그 주요 프로세스는 [그림 15]과 같다.



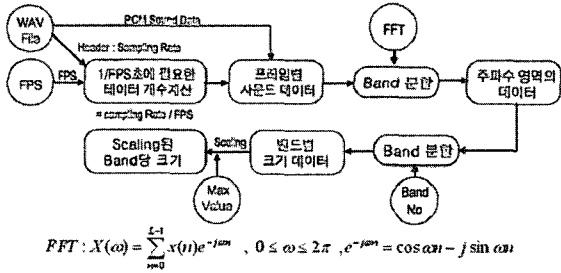
[그림 14] 음량 추출 주요프로세스

### 주파수 밴드 추출

기본적으로 ①WAVE 파일, ② FPS(Frame per Sec), ③최대값을 입력받는 것은 음량을 추출하는 것과 동일하며, 추가적으로 ④FFT Point, ⑤밴드의 숫자, ⑥Frequency Threshold, ⑦ 데시벨 사용 여부를 입력받는다.

우선 입력된 FPS(Frame per Sec), 즉 초당 프레임 수와 대상 웨이브파일의 샘플링 레이트를 기반으로 매 프레임(1/FPS초) 마다 필요한 사운드 데이터의 개수를 식(5)에 의해서 구한다. 사운드 데이터를 프레임마다 필요한 개수에 맞게 나누고 식 (3)을 사용하여 입력된 FFT Point에 맞게 FFT를 수행하여 주파수 영역에서의 데이터를 계산한다. 이 데이터를 가지고 각 주파수별 크기를 계산하는데 ⑦의 데시벨 사용 여부가 체크되어 있으면, 데시벨을 사용하여 크기를 계산한다. 크기에 데시벨을 사용하는 이유는 인간의 청감 특성에 더 적절히 대응하기 위해서이다. 하지만 데시벨을 사용하면 크기의 변화가 적어져서, 최종적인 시각화에도 그 역동성이 줄어들다는 단점이 있다. 다음으로 ⑤의 Band No에 따라서 주파수 밴드를 분할한다. 기본적으로 인간의 가청 주파수 대역인 20~22050Hz의 범위를 분할하며, 위에서 언급한 선형 밴드(Linear Band)와 로가리즘 밴드(Logarithms Band)를 동시에 구하도록 되어 있다. 여기서 ⑥Frequency Threshold의 값이 22050이 아니면 20~Frequency Threshold의 값을 Band No에 맞게 나눈다. 이는 Frequency Threshold 이상의 주파수는 버리고 계산하는 것으로, 인간의 청각이 고주파 보다는 저주파에서 감도가 높다는 사실에 기반하고 있다. 즉, 감도가 낮은 고주파쪽은 버리고 감도가 높은 저주파쪽을 집중적으로 분석하여 그 시각적 효과를 극대화시키기 위한 것이다. 이렇게 밴드가 분할되면, 각 밴드에 해당되는 주파수의 크기 값들을 모두 더해서 밴드별 크기를 구한다. 구해진 크기는 다시 ③의 최대값에 맞게 스케일링(Scaling) 되어, 최종적으로 각 프레임에 해당하는 밴드별 크기를 출력한다. 그 주요 프로세스는 [그림 15]와 같다.



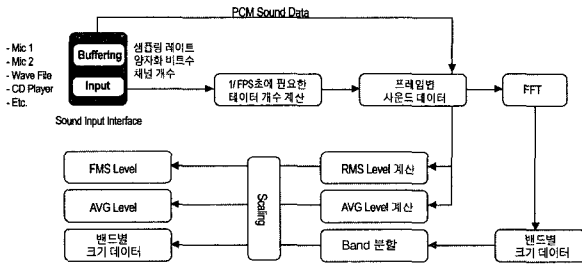


$$FFT: X(\omega) = \sum_{n=0}^{N-1} x(n)e^{-j\omega n}, \quad 0 \leq \omega \leq 2\pi, \quad e^{-j\omega n} = \cos \omega n - j \sin \omega n$$

[그림 15] 주파수 추출 주요프로세스

### 3-5. 실시간 구현 방법

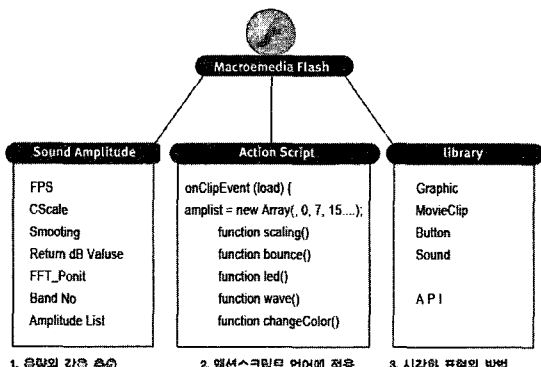
위에 구현된 프로그램은 실시간 사운드 데이터를 처리하는 것은 아니지만, 실시간 처리에서도 기본 로직은 그대로 활용할 수 있다. 입력으로 받는 웨이브 파일 대신 마이크 또는 다른 사운드 소스 장비로부터 받는 인터페이스(Interface)를 구현하고, 실시간으로 받는 데이터 처리를 위한 적정 수준의 버퍼링(Buffering)을 사용하면 된다. 물론, 사용하는 PC의 연산처리 능력에 맞게 FFT 포인트와 밴드의 숫자를 적절히 조절해서 사용하여야 한다. 또한, 여러 개의 사운드 소스를 입력으로 하고 동시에 프로세스를 병렬로 수행하여, 다수의 사운드 입력을 동시에 분석하여 시각적으로 표현하는데도 활용할 수 있다.



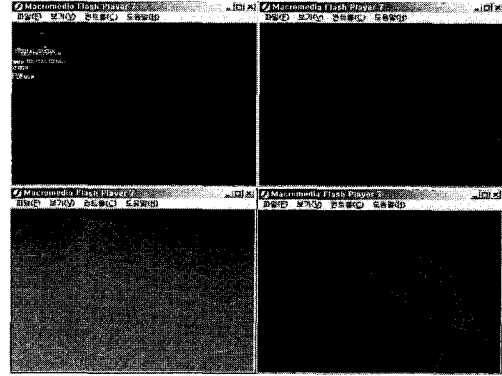
[그림 16] 실시간 분석 개념도

## 4. 인터랙티브 디자인 및 고찰

### 4-1. 플래시를 이용한 제작구현

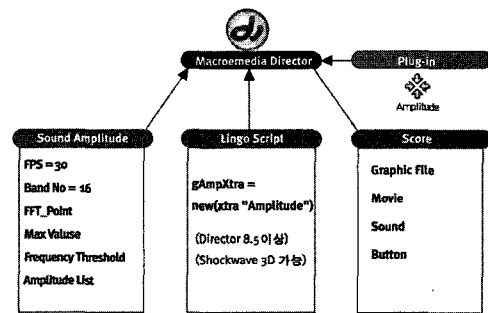


[그림 17] 추출된 주파수 및 음량의 값에 반응하는 플래시 구조



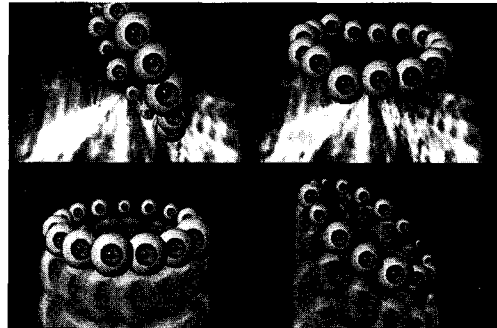
[그림 18] 충북과학대학교 졸업전시 발표, 2004

### 4-2. 디렉터를 이용한 제작구현



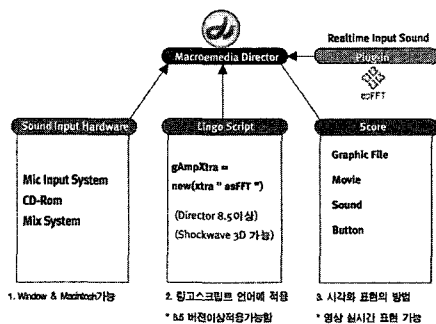
1. 음량의 값 추출  
• 밴드와 나눈값 적용
2. 링크스크립트 언어에 적용  
• 8.5 버전이상적용가능함
3. 시각화 표현의 방법  
• 영상 실시간 표현 가능

[그림 19] 추출한 음량 및 주파수 값에 의한 디렉터 구조

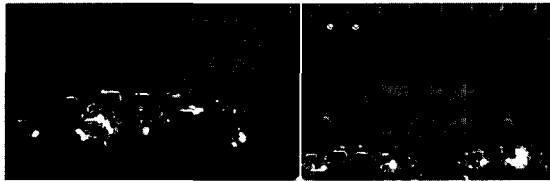
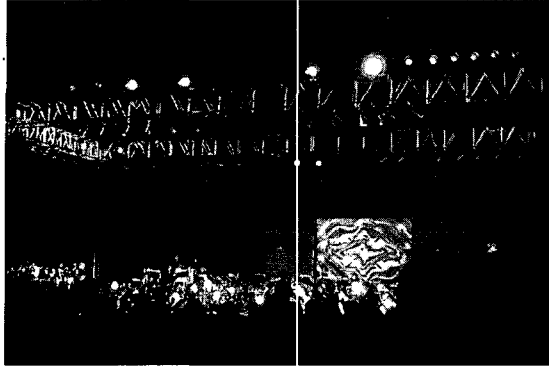


[그림 20] 중앙대학교 첨단영상제 발표, 2004

### 4-3. 디렉터를 이용한 리얼타임 제작구현



[그림 21] asFFT Xtras에 의한 실시간의 디렉터 구조



[그림 22] 신촌 아트 페스티벌 개막공연 발표, 2004

## 5. 결 론

본 논문에서 살펴 본 바와 같이 멀티미디어 콘텐츠에서 현상적으로 나타나는 속성의 하나인 상호작용(interaction)이라는 개념은 기술과 더불어 발전해 온 인터랙티브 디자인의 미술사적, 사회학적 배경을 통해 그 영역을 조금씩 넓혀 가고 있음을 알 수 있었다.

미디어를 둘러싼 대중문화의 지위행상과, 포스트 모더니즘의 시대에 주지하게 된, 주체와 타자와의 문제, 기술의 발전과 더불어 맞게 되는 예술작품 생산과 예술가의 개념에 대한 논의들 등의 그 이론적 배경으로 설명되었다. 사회 제반의 모든 면과 결부시킬 수 있을 정도로, 그 배경과 영향은 복잡한 양상을 띠고 있다는 것을 다시 확인하였다. 그만큼, 인터랙티브 디자인은 하나의 양식으로 현대사회에 퍼져 있었으며, 특히 그 중“상호작용”이라는 개념을 띠고 나타나는 작품들은, 더욱 괄목할 만한 발전을 하고 있다.

많은 창작가들이 이러한 기술력에 힘입어 일방적인 예술이 아닌 상호 소통하는 체계로서의 예술을 고안하고 개발하여, 이전에는 발견할 수 없던 새로운 감각과 자극을 만들어 내고 있다. 또한 그 감각을 통해 우리가 타인과 커뮤니케이션하기 위해 사용하는 모든 기법을 이용한다. 대부분의 경우 우리는 (글로 적어놓고 난 직후)먼저 그래픽 디자인, 비디오 예술, 영화 촬영술, 인쇄술, 삽화, 사진술과 같은 시각 디자인 분야를 이용할 것을 생각한다. 청각 디자인과 공학, 음악 공연 그리고 음성애 의한 커뮤니케이션도 적절한 상황에서는 매우 효과적이다. 실제로 그것들이 가장 창조적인 매체인 경우도 있다.

이러한 효과적인 쌍방향의 디자인(예술)이 우리시대에 더욱 긴밀하고 새로운 소통의 방법을 제시하고, 더 넓은 의미의 “상호소통”을 촉진시킬 수 있다면, 디자인(예술)이 갖는 본래의 기능들에 일조할 뿐 만 아니라, 세계를 보는 시각에 있어서도 더욱 다양한 방법과 비전을 제시할 수 있을 것이라고 본다.

## 참고문헌

- Clement Mok: Designing Business, (1996).
- 이만재. 이상선 공저: 멀티미디어 교과서, (1998).
- 김미지자: 감성공학, (1988).
- 이재린: 멀티미디어 환경에서 정보디자인 제작구성 요소에 관한 연구, 디자인과학연구, (2003).
- 안정미: 감각기관을 통한 정보의 상호간섭에 관한 연구 석사학위논문, (2000).
- 이중재: 웹아트에 있어 소통성의 의의와 사회적 역할 석사학위논문, (2001).
- 고용환: 음향기술이론과 응용, (1999).
- 박현재: 콘텐츠 산업과 디자인 벤처, 월간디자인, (1999).
- 정승남: 멀티미디어 인터페이스 디자인에서의 시각적 내러티브 기법의 적용에 관한 연구, 한국 과학 기술원 산업디자인학과 석사논문, (1999).
- 박영민: 웹 그래픽 인터페이스의 시지각 평가에 대한 연구, 한국과학기술원 산업디자인학과 석사논문, (2000).
- Robert Jacobson: Information Design, 안그라픽스, (2002).
- 홍원기: Strategic Interactive Communication Design Research for the Multimedia Contents, 경원대학교, (1999).
- 김하진, 이만재, 권은숙, 교육: 디지털 콘텐츠, 안그라픽스, (1999).
- 윤준수: 인터넷과 커뮤니케이션의 패러다임의 대전환, 커뮤니케이션북스, (1998).
- Painterly Interfaces for Audiovisual Performance, Golan Levin, B.S. Art and Design Massachusetts Institute of Technology May, (1994).
- 김하진, 이만재, 권은숙, 교육: 디지털 콘텐츠, 안그라픽스, (1999).
- 한학용, 하성욱, 허강인: Multimedia Sound Programming. 영진출판사, (2003).
- 김경민, 김수정, 김태균, 이성식, 이종호, 장동훈, 한민형, 홍석기: The Web, 이디자인, (2002).
- 제프 래스킨지음., 이진표 옮김: Humane Interface, 안그라픽스, (2003).