

# 모션 트래킹의 센싱화 작업을 이용한 인터랙티브 아트

이준의<sup>1</sup>, 배성준<sup>2</sup>, 김형기<sup>3</sup>  
중앙대학교 첨단영상대학원 예술공학 연구실  
reejun@nate.com<sup>1</sup>, email@imagelab.cau.ac.kr<sup>2</sup>, unzi@cau.ac.kr<sup>3</sup>

## Interactive Art using a Sensing task of Motion Tracking

Lee Jun Eui<sup>1</sup>, Bae Seong Joon<sup>2</sup>, Kim Hyeong Gi. Prof<sup>3</sup>  
Digital Art Lab, Graduate School of Advanced Imaging Science,  
Multimedia & Film, Chung-Ang University

### 요 약

탈 장르화와 영상의 다양화로 예술에서의 표현의 한계는 극대화 되고 있으며, 디지털 매체를 통한 인터랙션 역시 디지털 아트에서의 보편적인 표현 방법으로 전환 되었다. 예술에서의 상호작용은 보여주는 것에서 참여하는 것으로의 전환을 꾀 하고 있고, 디지털을 기반으로 한 다매체, 다중화가 이를 뒷받침 한다.

상호작용을 위한 디지털 아트 작품을 위해서는 센서를 이용한 표현방법들이 있으나, 기타의 센서도구를 사용하지 않고, 단지 카메라로부터 입력된 신호만을 기반으로 영상과 관객의 상호작용을 끌어 낼 수 있다.

이를 위해서는 모션트래킹을 위한 알고리즘을 응용함으로써 사물의 밝기 값에 충족한 데이터를 만들어 낼 수 있고 그에 해당하는 밝기의 변환 값을 이용해 미세한 사물의 변화를 감지 할 수 있는 것이다.

이런 일련의 영상작업을 위해서는 사물의 움직임과 반복성을 얼마만큼 인지하고 감지 해내느냐고 관건인데, 다시 말해 한번 측정한 움직임이 있는 사물의 데이터도 프로그램상에서 계속하여 인지가 되어야 하고, 그 데이터 값을 영상으로 반환하여야 하는데, 이는 영상의 지속적인 변화를 가져 올 수 있다는 걸 의미한다.

따라서 본 논문에서는 모션 트래킹의 기본 알고리즘을 제시하고 영상작품과 인터랙션작품의 변환을 위해 사용된, 센서 대체도구인 웹 캠의 데이터 즉, 색상 값과 밝기 값을 적절하게 활용함으로써 표현의 다양성을 이끌어 내고 디지털 인터랙티브 작품으로써의 제작을 꾀하고자 한다.

**Keyword:** Motion Tracking, Interactive Art, Digital Art, Sence

### 1. 서 론

디지털 미디어 아트는 빠른 기술의 진화로 인해 많은 변화를 가져왔고, 그로 인한 작품에서의 표현방식은 상호작용이라는 실시간 디지털 처리 방법을 택함으로써 다양한 멀티미디어를 이용한 기술의 융합(convergence)을 이루었다.

센서를 이용한 실시간 인터랙티브 작품에서는 모든 행위예술과, 또한 예술가가 관객의 참여를

기대하는 작업과 작품을 만들어 낼 때 대단히 중요한 역할을 하는 개념이다. 컴퓨터나 멀티미디어 시스템이 제작된 알고리즘이나 소프트웨어에 따라 관객의 반응을 매우 짧은 시간 동안 반응되어 보여지는 작업의 시간들인 것이다.

여기서 말하는 센서라는 것은 일반적인 정보처리 프로그램에서 사용되는 물리적 자료를 말한다. 제스처, 신체, 시각, 목소리 등을 포착하고

전송하는 역할을 하는 기술장치이다. 이러한 센서는 40 여전 오토메이션(Automation: Automatic operation 의 약자)에 관한 전문서적에 오늘날의 센서 대신 트랜스듀서(transducer)라는 용어가 등장 하였고, 여기서 트랜스듀서란 변환기(變換機)를 의미한다. 한마디로 센서라 부르지만 그 범위는 매우 넓어 한마디로 표현하는 것은 매우 곤란하다, 그러나 조금 극단적으로 표현하면 센서란 “어떤 정보 및 에너지의 검출수단이다”라고 정의 할 수 있다.

본 논문에서는 앞서 정의한 일반적인 전기신호를 얻기 위한 물리적, 화학적, 생물학적 수단으로 전자공학에서의 회로를 사용한 센싱 작업과는 구별된다. PC 용 웹 카메라로 입력된 영상신호를 기반으로 인간의 눈을 대신하여 사물의 형태를 감지하는 센싱화 작업을 통한 지능형 센서 즉 비전센싱 작업을 제안 하였다. 컴퓨터 디지털 기술에 기반을 둔 실시간 모션트래킹의 센싱화 알고리즘은 다양한 작품과 여러 형태의 인터랙션 된 영상을 관객에 보여줄 수 있는 공간연출을 가능하게 할 것이다.

## 2. 센싱기술과 비전 시스템, 모션트래킹

### 2-1. 센싱 기술

최근의 센싱 기술은 유비쿼터스 컴퓨팅을 위한 핵심 기술 중에 하나이며, RFID 칩을 이용한 무선 주파수 인식에서부터 우주기기, 가전제품, 그리고 아파트에 이르기까지 광범위한 분야에서 사용되는 중요한 기술이다.

학술적으로 센싱 기술이란 센서 디바이스와 신호처리 장치 등의 하드웨어와 신호처리 알고리즘 등의 소프트웨어로 이루어지는 시스템 즉 “센싱 시스템을 실현하는 기술”이다. 또 여기서 센싱 시스템이란 센싱 대상의 정보를 수집, 처리하여 유용한 정보를 제공하는 기능을 갖는 시스템”으로 정의되는 시스템이다.

이 센싱 시스템이 보다 정확한 것이 아니면 즉 센싱에 의해서 수집되는 정보가 바르지 못하면 기계나 시스템이 아무리 뛰어나다고 해도 그 기능과

성능을 충분히 발휘할 수가 없다

### <센싱을 위한 요소기술과 주요제품>

요소기술	주요제품
반도체	압력센서, 가속도센서, 인공망막칩, 포토다이오드
적외선	적외선카메라, 관측위성탑재용 센서
광. 레이저	차간거리계, 레이저토폴러, 진동계, 형상검사장치, 레이저식 변위계, 광엔코더
전자파	합성개구 레이더, 기상관측 레이더, 레이더식 속도계
방사선	원자력발전소용 방사선 모니터, 의료용선량계
초음파	초음파탐상기, 초음파두께계
전자기	전자식전력량계, 인텔리전트 전자유량계
화상처리	교통류계측장치, 투입감시장치, 지문조회장치, 로봇비전, 차간거리센서
시스템	송변전기기용고장진단시스템, 수배전전기기감시장치, 플랜트감시제어장치

<표 1>

### 2-2. 비전 시스템

컴퓨터 비전 또는 머신 비전이란 여러 매체 (Voltage, Ampere, Frequency, 시간, 온도, 위치 데이터)중에서 카메라를 이용하여 2 차원의 위치 데이터를 입력 받아 검사, 제어하는 방법으로서 비 접촉 방식이며, 인간의 육안 검사와 비슷한 역할을 하는 장치를 말한다.

비전 기술의 목표 : 화상처리를 근거로 하여 인간의 시각 능력에 유사한 지적 기계를 만들어 가고자 한다.

- 비전 기술 : 컴퓨터 비전, 화상처리, 머신 비전, 로봇비전 등

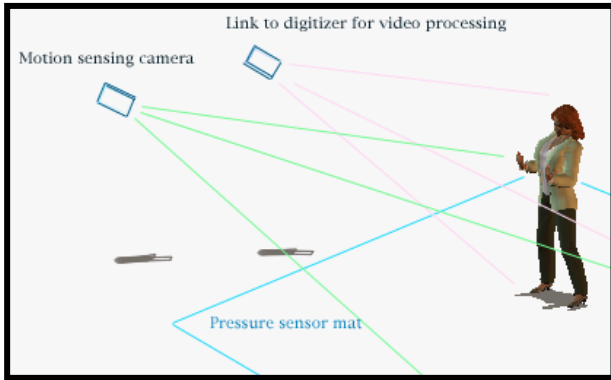
- 컴퓨터 비전 : 일반적으로 컴퓨터에 의하여 만든 계산을 하여, 화상 처리를 활용하고자 하는 분야

- 화상처리 : 기존의 1 차원적인 전기적 신호처리에 대비하여, 2 차원적인 정보에 대한 다양한 컴퓨터 또는 회로적 처리를 수행하는 기술분야를 총칭

- 머신 비전 : 기계에 시각을 부여하고자 하는 분야

### 2-3. 비전 시스템의 구성

비전시스템의 구성은 조명과 렌즈, 카메라, Frame Grabber 으로 이루어 진다.



<그림 1>

- 조명 : 화상을 입력 받기 위해 검사하고자 하는 부분의 화상을 선명하게 얻기 위하여 검사부분이 명확하게 들어날 수 있도록 조명을 만들어 주는 역할을 한다
- 렌즈 : 카메라의 센서 부분에 빛을 모아주는 장치로서 대상물의 크기, 거리와 관계가 있다
- 카메라 : 이미지를 촬영하는 부분(Image Acquisition Unit)과 비디오 신호를 출력하는 부분(Image Output Unit)으로 나눌 수 있다. Image Acquisition Unit 은 CCD(Charge Coupled Device)소자로 구성되는데, 이것은 픽셀(Pixel)이라 불리는 개별의 빛에 감응하는 센서들로 이루어져 있다
- Frame Grabber : Frame Grabber 는 영상 입력 장치에서 들어오는 영상 신호를 디지털화 한 후 내장된 메모리에 영상을 기억하거나 실시간으로 영상을 출력하는 장치이다.

### 2-4. 모션 트래킹(Motion Tracking)

영상처리 기술의 한 분야인 윤곽선 검출과 움직임 추적은 그 응용 범위가 넓다. 또한 이 분야는 컴퓨터 비전에서 고전적인 관심분야이고, 현재에도 많은 연구가 이루어 지고 있는 실정이다. 현재 대부분의 윤곽선 검출과 움직임 추적의 기술은 의학과 보안 시스템, 3 차원 객체를 추출, 추적 및 인식하여 재구성하는 컴퓨터 비전, 실제 환경에서 컴퓨터 그래픽 영상을 추출, 추적하여 실제와 가

상 영상이 혼합하는 증강현실, MPEG-4, MPEG-7 등과 같은 객체 기반 비디오 압축, 애니메이션을 위한 움직임의 특징을 정확히 표현하기 위한 행동양식 분석(HCI:Human Computer Interaction), 홈네트워크등 많은 응용분야에서 사용되고 있다. 감시하는 장소에서 움직임이 검출되면 그 검출된 물체의 윤곽(Edge)를 모니터 상에 Display 시키고, 그 움직임을 추적하는 기능을 제공한다



<그림 2>

일반센서의 물리적 한계로 인한 대안으로 사용된 모션 트래킹은 센서와 유사한 결과를 얻을 수 있지만 회로를 이용한 센서나, 무선 센서 그리고 CATV 카메라, CCD 카메라, 아주 저렴한 가격의 웹 캠(PC CAMERA)으로 카메라 렌즈의 넓은 화각을 이용하여 보다 확장된 공간 영역을 동시에 통제할 수 있다는 점을 장점으로 들 수 있다. 또한 사물의 움직임, 전시 공간에서의 관객의 움직임 전체 뿐 만 아니라 사람의 얼굴을 트래킹해서 얼굴의 각도나 위치에 따라 제스처마다 구분된 영상의 입력 값을 받아 들일 수 있다.

### 2-5. 센싱작업을 이용한 인터랙티브 디지털 미디어 아트

센서 작업을 이용한 인터랙션 작품으로는 칼 심스(Karl Sims)의 <갈라파고스 Galapagos>에서 찾아 볼 수 있다.

이 작품은 열두 대의 연결된 모니터로 구성된다. 열두 대의 텔레비전 모니터는 제각기 관객이 발로 작동시킬 수 있도록 센서가 달려있고, 관객이 각

각 모니터 바닥에 설치된 센서를 발로 누르게 되면 그 발판에 위치한 모니터 화면의 이미지는 점점 커지면서 화면 전체에 가득 차게 된다. 준비된 영상은 제각기 다른 종류의 화면으로 흑백이거나 컬러, 식물, 동물, 뱀, 등. 관객이 선택한 ‘생명체’는 유희적 방식으로 형태가 만들어지거나, 커지거나, 변형된다. 관객은 이러한 시각적 유희에 큰 즐거움을 느낄 수 있다.

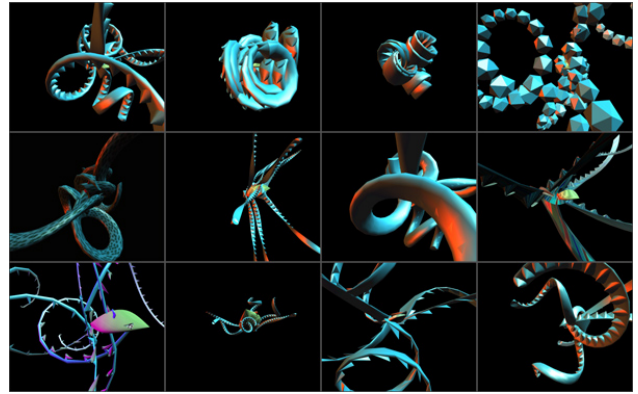


.<그림 3>



< 그림 3,4 > 칼심스 <칼라파고스> 1997

미국의 컴퓨터 과학자 칼 심스는 진화적 예술의 대표주자이다. 심스는 인공진화 시뮬레이터에서 컴퓨터상의 3 차원 세계에 여러 블록으로 이루어진 가상생물을 만들었다. 가상생물은 센서입력을 하면 방향그래프에 담긴 유전정보에 따라 적절히 행동한다. 블록을 제어하는 신경망은 센서입력에 대해 움직임을 만드는 자극 반응 계로 활동한다. 예컨대 ‘걷는다’ ‘헤엄친다’ 등의 과제를 주면 서로 짝짓기와 돌연변이 등의 작용을 벌인다. 그런 과정을 통해 태어난 가상생물은 다채로운 형상과 행동을 보여준다



<그림 5> 칼심스

또 다른 센서를 이용한 인터랙티브 멀티미디어 작품으로 두젠준(Du Zhenjun)의 <두 젠준 박사의 해부학 강의>라는 작품을 예로 들 수 있다.

관객은 커다란 화면과 마주친다. 실제 나무로 만든 수술대가 절반은 화면 위에 놓인 채로 관객 쪽을 향하고 있는가 하면, 나머지 절반은 이미지로 된 가상현실로 뻗어 있다.

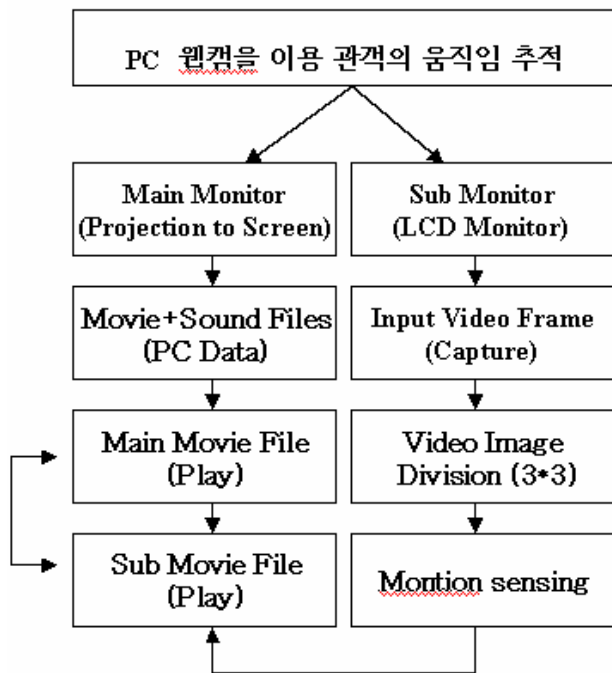


<그림 6> 두 젠준 <두 젠준 박사의 해부학 강의>

뒤편에 서 있는 여덟 명의 두 박사들이 지켜 보는 가운데 해부학 실험을 기다리고 있다. 센서는 방문객의 실제 접근을 감지하고, 방문객이 화면 가까이에 오면 여덟 명의 두 박사 분신들이 화면 저쪽에서 나타나 관객 앞에 자리를 잡는다. 관객이 그 곳을 떠나면 여덟 명의 박사들도 해부학 실험실을 떠나 가상현실 속으로 돌아가는 작품이다.

### 3. 제안된 모션트래킹 알고리즘의 센싱화

본 논문에서는 디지털 인터랙티브 아트작품을 위해 제안된 모션 트래킹의 알고리즘의 프로그램 구성도를 <그림 7>에 나타내었다.



<그림 7> 모션트래킹의 센싱화 프로그램 구성도

#### 3-1. 작품구현을 위한 모션 트래킹 영상

회로가 적용된 센서를 사용하지 않고 보편적인 웹 캠을 사용한 센싱화 작업은 범용적 작품 구현이 가능하다. 이를 이용한 인터랙티브 미디어 작품의 모션 트래킹(움직임 추적)은 영상 입력 부분에 필수적으로 구현되어야 영상의 변화를 감지할 수 있다.

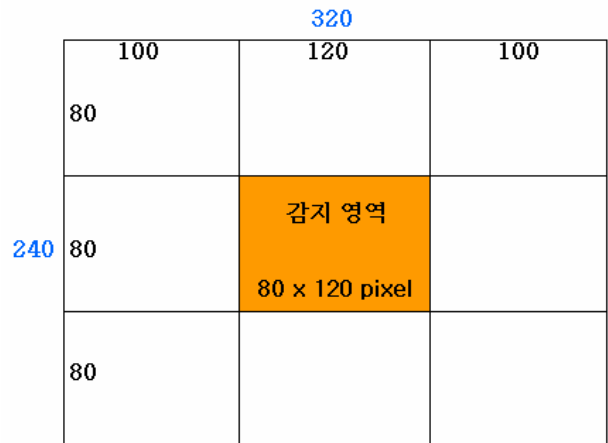
일반적인 회로로 제작된 센서 시스템은 거리와 반응 거리가 고정 되어 있고, 복잡한 배선과 한정된 행동반경의 협소함이 단점이다. 이런 센서 시스템의 단점을 보완한 객체 추적(Motion Tracking)은 센서 시스템과 같은 결과를 얻을 수 있지만 대중적으로 사용하는 간단한 웹 캠만을 이용하여 임의의 지정된 일정 공간을 감지할 수

있고, 센서 시스템 보다 광범위한 지역을 동시에 통제할 수 있다.

객체 추적 즉 모션트래킹의 알고리즘을 이용하기 위해서는 윈도우 운영체제 기반의 Micro Soft 사의 프로그램 제작 툴인 Visual C++로 구현하게 된다.

물론 기존의 상용화된 제작 툴들이 많이 사용되고 있긴 하지만 초기 구입비용과 사용상의 한계가 뒤따른다. (MAX/MSP, JITTER, NATO, 등)

이렇게 제작된 프로그램 소스는 다양한 작품의 변형과 적용이 가능하고 표현의 확장이 용이해서 사람의 얼굴을 트래킹 한다면 얼굴의 각도나 위치에 따라 변화를 줄 수 있고, 관객의 제스처나 위치 마다 다르게 인식될 수 있어 쌍방향 인터랙션 작품 제작이 가능 해 진다.



<그림 8> 입력 영상의 분할 영역

PC USB 용 웹 캠으로부터 입력된 영상은 각 이미지 별로 총 9 개의 분할된 영상으로 인식이 되고, 분할된 9 개의 부분 중 임의의 한 영역을 지정한 다음 그 부분의 움직임이 감지되면, 미리 제작된 동영상과 사운드가 반응하여 재생하게 된다. <그림 8>

이때 입력된 영상의 사이즈는 320 x 240 사이즈이며, 그 중 가로 3 줄, 세로 3 줄 총 9 개의 분할된 영역 중 가운데 영역의 해당 부분이 바뀐 픽셀의 수가 일정 수 이상일 때와 바뀐 픽셀이 얼마나 많이 변했는가 하는 두 가지 조건으로 감지 한다.



### 3-2. 제안된 알고리즘의 구현

일반적으로 사용하는 웹 카메라는 사용자의 얼굴을 비추는 것에 맞추어져 있기 때문에 사용자의 전신을 감지해야 하는 본 논문에서 사용하기 위해서는 카메라에 입력되는 영상의 일부만을 사용해야 하므로 전체 영상 중 중심의 일부 화소(M by N)만을 사용하였다. 본 논문에서는 영상의 일부만을 활용했지만, 제작자의 의도에 따라 사용자의 위치 별로 다른 상호작용이 가능한 시스템(작품)을 구성할 수도 있다

사용자(관람자)의 움직임이 파악하기 위해서는 카메라로 입력되는 영상의 변화를 감지해야 한다. 따라서 <수식 1>과 같이 현재 영상과 이전 영상 화소의 명암 값을 비교해서 그 차이가 일정한 한계치(threshold) 이상이면 해당 화소의 변화가 있는 것으로 감지하도록 하였다. 여기서  $f(x,y)$ 는 현재 영상의 명암 값,  $F(x,y)$ 는 이전 영상의 명암 값이고,  $(x,y)$ 는 영상의 화소가 위치한 좌표 값이다.

$$\begin{aligned} & \text{if } |f(x,y) - F(x,y)| > 30, P(x,y) = 1 \\ & \text{else } P(x,y) = 0 \end{aligned}$$

<수식 1>

수식(1)과 같이 영상의 각 화소별 변화를 검사하여 <수식 2>와 같이 변화된 화소의 총 합계를 구하게 된다

$$T = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} P(x,y)$$

<수식 2>

변화된 화소의 총 합계가 전체 화소의 일정비율 이상일 경우 사용자의 움직임이 감지된 것으로 판단하고 영상을 바꾸게 된다<수식 3> 여기서 경계 값(V)는 0 과 1 사이의 값으로 사용자에게 의해 조절이 가능하며, 0 에 가까워 질수록 사용자(관람자)의 작은 움직임에도 미세하게 반응한다.

$\text{if } T \times V > M \times N$ , 영상 변화  
 $\text{else}$  현재 영상 유지  
 <수식 3>

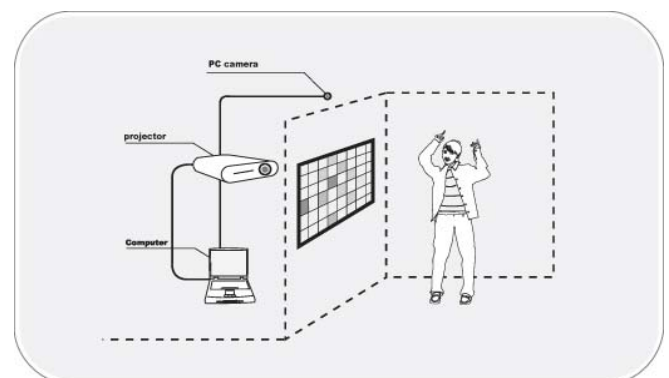
### 3-3. 작품 영상 소스

이렇게 제작된 프로그램은 영상 입력을 기반으로 웹 카메라로 입력된 지정된 영역 신호에 따라 미리 제작된 메인 영상과 서브 영상을 재생하게 된다.

관객의 움직임이 감지 되기 전에는 PC 에 저장된 메인 영상이 스크린에 재생되고, 웹 카메라의 일정 지정영역에 움직임이 감지되면, PC 에 저장된 두 번째 영상이 관객의 움직임에 의해 스크린에 프로젝션 된다.

관객의 움직임에 반응하여 재생되는 영상의 재생 시간은 동영상이라는 특성을 감안하여 메인 영상과 서브 영상의 재생 시간을 같게 제작하였다. 관객의 움직임에 반응하여 인터랙션 되는 영상을 프로젝션 하기 위해서는 스크린에 전체영상으로 재생하여야 한다,

따라서 스크린용 프로젝터 외에 컨트롤용 서브 모니터가 필요하고, 서브모니터에서는 프로그램을 실행하여 관객의 움직임을 트래킹 하며, 반응에 따라 프로젝트를 통해 메인 영상과 서브 영상이 교차되어 각기 개별적인 사운드와 함께 재생된다.



<그림 9> 작품설치를 위한 시스템 구성

#### 4. 실시간 인터랙티브 작품 구현

본 논문에서 제안한 방법으로 구현된 실시간 인터랙티브 작품은 기존의 다양한 산업용, 공학용, 계측용 센서를 웹 캠으로 대체하고 객체 추적의 센싱화 작업을 근간으로 하여 제작 되었다.

##### 4-1. 작품 컨셉 (리듬과 흐름)

평온한 바다 속 흐름을 하나의 리듬으로 표현한 작품.

물의 흐름과 물 고기떼의 흐름이 마치 하나가 되어, 음악에 맞춰 춤을 추듯 리듬을 갖는다. 때로는 느리고 때로는 빠르게..

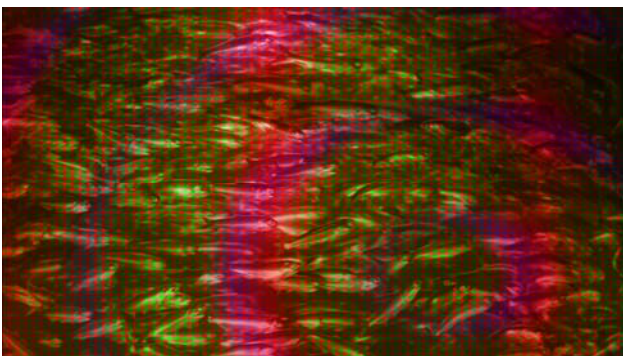
현실 속 관객은 그 영상 속으로 빠져든다. 리듬에 동화된 관객은 마치 바다 속 흐름에 맞춰 움직이는 리듬과도 같다.

이것은 바다만이 들려주는 현실 속의 리듬이 아닐까...?

관객의 움직임에 따라 그 리듬은 빠른 흐름으로 또는 느린 흐름의 영상과 사운드로 인터랙션 된다

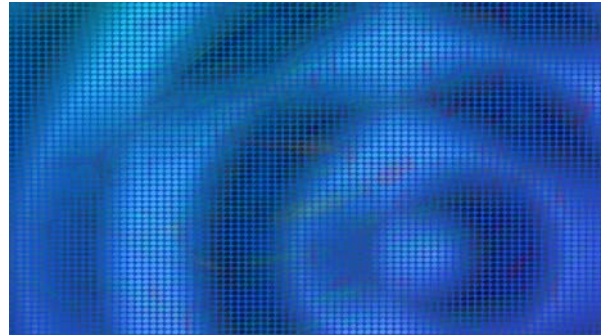
인간이 인간과 유기적인 관계를 맺고 살아 갈 수밖에 없는 그래서 인간의 움직임에 반응할 수밖에 없는 모습을 바다의 리듬과 흐름을 통해 표현하고자 하였다

<인터랙션 前 잔잔한 흐름의 바다>



< 그림 10> 리듬과 흐름, 이준의, 2005

#### <관객의 움직임에 반응한 영상>

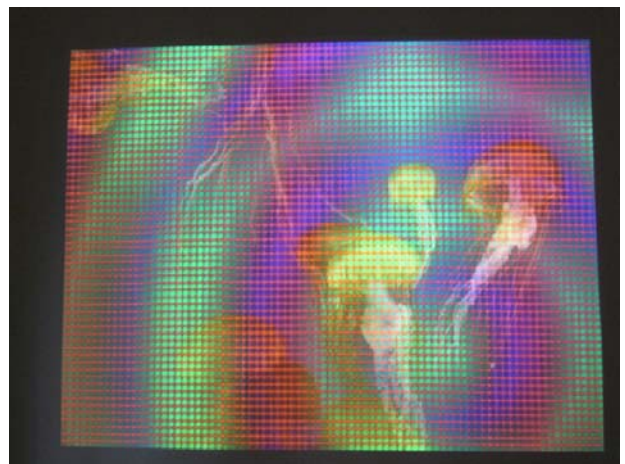


< 그림 10> 리듬과 흐름, 이준의, 2005



<그림 11> 바다 네 품에 안기다, 라메르, 2005

#### <2006 의정부 국제디지털미디어아트 수상작>



<그림 12> 리듬과 흐름, 의정부 예술의전당, 2005

## 5. 결론

디지털이라는 매개체를 이용한 디지털 미디어 아트는 예술가에게 더 이상 아날로그적 표현을 요구 하지 않은 듯 하다. 수 많은 작가들이 새로운 영상과 표현을 위해 다양한 방법으로 관객에 대한 접근을 시도하고 있고, 그 접근 방법으로 상호작용 즉 인터랙티브라는 쌍방향성을 이용함으로써 관객은 더 이상 눈으로 보여지는 세계 속에 머물러 있지는 않다.

관객의 몸짓과 표정, 반응 하나 하나의 움직임을 놓치지 않고 작품 속에 몰입하게 하기 위해서는 디지털 시스템을 기반으로 한 센싱 작업이 유일한 방법이었으나, 비전 시스템의 모션 트래킹을 보안이나, 감시를 위한 도구로써 뿐만 아니라 실시간 인터랙티브 예술 작품을 위해 제작 구현 하였다. 이러한 실시간 모션 트래킹의 알고리즘 구현을 위해서는 디지털 영상 처리기술이 반드시 필요하며 프로그램을 제작을 위한 프로그래밍 언어에 대한 이해도 필요할 것이다. 이제는 과학이나 공학을 더 이상 특정한 분야의 도구로써가 아니라 예술작품 제작의 표현도구로서 발전 시킬 수 있는 모션 트래킹의 센싱화 작업을 제안하였고 그 예로써 작품을 제작하여 전시 하였다.

향후 연구에서는 웹 카메라를 통하여 입력된 각 영역별 분할된 픽셀위치에 따라 다양한 영상이 랜덤 하게 관객의 움직임에 반응하여 상영 될 수 있도록 할 것이다.

## 참고 문헌

- [1] 곡요흔사, 센서의 기초, 대영사, 2001
- [2] 신정황, 센서, 한진, 2001

[3] 플로랑스 드 메르디, 예술과 뉴테크놀러지, 열화당, 2005

[4] 오창근, 디지털 인터랙티브 미디어 공간 연출의 방법론, 한국영상학회 논문집 2호, 2002

[5] 김상진, 실시간 감시 시스템을 위한 사전 무학습 능동 특징점 모델 기반 객체 추적, 중앙대학교 첨단영상대학원 석사학위 논문, 2004

[6] Liang Wang, Weiming Hu, Tieniu Tan, Recent Developments in Human Motion Analysis, National Laboratory of Pattern Recognition Institute of Automation, Chinese Academy of Sciences, Beijing, P. R. China, 100080, 2002