

3D 센서테이블에서의 맨손을 이용한 컴퓨터와의 상호작용

이재선¹, 박경신²
한국정보통신대학교 디지털미디어연구소^{1,2}
{rosierr¹, park²}@icu.ac.kr

Bare-hand Human Computer Interaction on the 3D sensor table

Jae Seon Lee¹, Kyoung Shin Park²
Digital media lab., Information & Communications University^{1,2}

요약

3D 센서테이블은 디스플레이가 내장된 테이블 위에서 사용자의 맨손이 움직일 때 그것을 감지하고 반응하는 센싱정보를 처리하도록 만든 것으로 양안시차 방식의 3 차원 입체영상에서 보여지는 가상의 오브젝트와 사람과의 인터랙션이 가능하도록 설계되었다. 본 연구에서는 3 차원 공간에서의 맨손을 이용한 사람과 컴퓨터 간의 인터랙션 방법에 중점을 두고 있는데 맨손의 위치 정보와 제스처를 알기 위해서 Electric Field Sensing 기법을 이용하였다. 현재 구현된 3D 센서테이블은 테이블로부터 30cm 이내의 거리에서 움직이는 맨손의 위치와 간단한 제스처를 인식할 수 있다. 이를 통해서 기존의 마우스나 키보드가 가졌던 3 차원 공간에서 사용할 때 불편했던 점들을 해소하고, 직관적이면서도 누구나 사용하기에 쉬운 인터페이스 방법을 제시함으로써 사람과 컴퓨터와의 인터랙션 방법을 확장시킬 수 있다. 또한 기존의 많은 연구에서 맨손을 이용한 기법들이 주로 2D 기반이었던 것에 비하여 본 시스템은 2D 뿐 아니라 3 차원 공간에서도 활용할 수 있다는 장점을 가진다.

Keyword : Bare-Hand, Human Computer Interaction, Electric Field Sensing

1. 서론

우리가 흔히 사용하는 컴퓨터의 경우에 입력장치로서 키보드와 마우스가 보편적이다. 이 두 가지 장치는 2D 기반의 애플리케이션에서 최적화된 인터페이스를 제공하고 있으며 대부분의 사람들이 약간의 학습으로 컴퓨터를 쉽게 익히고 사용할 수 있게 해준다. 그런데, 3 차원 애플리케이션의 경우에는 마우스나 키보드가 사용자에게 적합한 인터페이스가 되지 못하는 것이 사실이다. 왜냐하면 마우스나 키보드는 2D 상의 위치정보를 기반으로 설계된 것이기 때문에 3 차원 애플리케이션에서 사용되는 3D 위치정보를 나타내기가 힘들기 때문이다. 그렇기 때문에 3 차원 애플리케이션을 위해서는 기존의 입력장치와는 다른 3 차원 위치정보

를 기반으로 하는 새로운 인터페이스가 필요하다. 본 연구는 3 차원 입체영상을 보여주는 3D 디스플레이에서 사용할 수 있는 3D 인터페이스를 개발하는 것에서 시작되었으며 특히, 맨손을 이용한 인터페이스 개발에 중점을 두고 있다. 본 연구의 목적은 3 차원 디스플레이에서 맨손을 이용하여 사람과 컴퓨터와의 인터랙션을 가능하게 하는 인터페이스를 개발하는 것이다. 3 차원 디스플레이 상의 30cm 이내의 공간에서 맨손의 움직임에 따른 위치정보 및 간단한 손 모양을 인식하기 위해서 Electric Field Sensing 기법을 사용했다.

3 차원 애플리케이션에서 사용되는 인터페이스는 지금까지 여러 가지 방법들이 제시되어왔다. 이러한 인터페이스는 3 차원 공간상에서의 위치정

보를 얻는 것이 중요한데, 보통의 경우에 사람의 몸에 센서를 달거나 센서가 부착된 도구들을 사용한다. 그러나, 이러한 것들은 사람에게 자연스런 움직임을 보장해주지 못하며 또한 사용하기에 앞서 학습이 필요하다는 단점이 있다. 그래서, 본 연구에서는 사람과 사람 사이에서처럼 자연스럽고 친숙한 인터랙션 방법을 이용하기 위해서 맨손을 이용한 방법을 찾게 되었는데 가장 많은 연구가 진행되어 온 것은 비전을 이용한 방식이었다 [1][2][3]. 이 방식은 하나 이상의 카메라를 이용하여 사용자의 움직임을 감지하는 것인데 보통의 경우에 2D 애플리케이션에서 사용되었으며 다음과 같은 문제점들을 갖고 있다. 첫째, 카메라의 frame rate 보다 빨리 위치정보를 얻기가 힘들다. 둘째, 영상처리를 할 때 입력대역폭과 계산 량이 영상의 크기에 비례하여 커지게 된다. 셋째, 배경화면이나 빛의 밝기와 같은 주변환경에 영향을 많이 받는다 [4]. 이러한 문제점들로 인하여 본 연구에서는 비전 방식이 아닌 Electric Field Sensing 기법을 이용한 방식을 채택하였다. 비전방식을 이용한 시스템에 비해 Electric Field Sensing 기법을 이용한 시스템이 갖는 장점은 다음과 같다. 첫째, 간단하면서도 저렴하게 시스템을 구성할 수 있다. 둘째, 정밀도나 화면의 크기를 크게 하기 위해서 비전방식과는 달리 트랜스시버의 수를 늘리기만 하면 된다 [6]. Rekimoto 가 제안했던 SmartSkin[5]이 본 연구에서 사용되는 Electric Field Sensing 기법을 이용하여 맨손의 움직임을 감지하는 시스템의 대표적인 사례이다. 그런데, SmartSkin 의 경우에는 2D 기반의 애플리케이션에 사용하기 위한 시스템이었기 때문에 손의 움직임을 감지할 수 있는 proximity 가 5~10cm 에 불과했기 때문에 3 차원 애플리케이션에 사용하기에는 부적합했다. 그래서, 본 연구에서는 기존의 센싱방법을 보완하여 약 30cm 이내의 공간에서 맨손의 움직임을 감지할 수 있도록 개선하였다.

3. 본 론

3-1 하드웨어 구조

3D Sensor Table 시스템에서는 맨손의 움직임을 센

싱하여 위치정보를 얻기 위해서 Signal processing unit 과 meshed transparent plate 를 사용한다. 다음의 그림은 Hardware design 에서의 개요도이다.

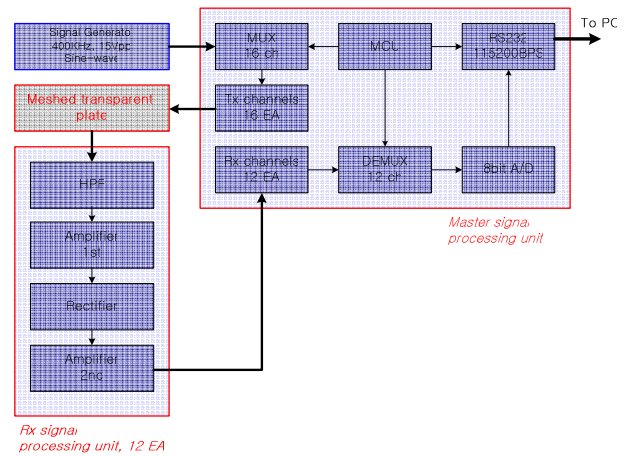


Figure 3-1 Overview of Hardware design

송신에 사용되는 신호는 400KHz, 15Vpp 의 Sine-wave 이다. 이 신호를 Master signal processing unit 의 MUX 를 통해서 16 개의 송신채널에 차례대로 전송하면, 이 신호가 Meshed transparent plate 를 거쳐서 Rx signal processing unit 의 12 개의 수신채널에 전송이 된다. 이 때 수신된 신호는 잡음이 섞여 있는 미약한 신호이므로 Filter 를 사용하여 400KHz 성분만을 뽑아내고 OP amp 로 증폭 및 정류하여 A/D 변환에 알맞은 신호로 조정이 된다. 그리고 나서 이 신호를 다시 Master signal processing unit 으로 보내면 DEMUX 를 거쳐서 A/D 변환을 거친 10 비트의 데이터가 115200BPS 의 속도로 RS232 통신을 사용하여 PC 로 전송이 된다.

3-1-1 Signal processing unit

Signal processing unit 은 2 부분으로 나눌 수 있는데 송/수신 채널을 제어하고 수신된 신호를 A/D 변환을 거쳐 PC 로 전송하는 역할을 하는 Master processing unit 과 수신회로간의 간섭을 막기 위해서 독립적으로 만들어진 수신신호를 처리하는 12 개의 Rx signal processing unit 으로 구성되어 있다.

Master processing unit 은 미국 Microchip 사의 PIC micro-controller 인 PIC16F877A 를 사용한다. Master processing unit 의 역할은 크게 4 가지로 나눌 수 있다. 첫째, Signal Generator 로부터 입력되는 신호를 16 개의 송신채널을 통해서 차례대로 전송하는

것이다. 16 개의 송신채널 중에서 한번에 하나의 채널만이 열리게 되므로 수신채널에서는 한 채널에 대해서 12 개의 수신 값을 얻게 되고 모두 16 개의 송신채널이므로 결국 한 frame 에 대해서 192 개의 수신 값을 얻게 되는 것이다. 그런데, 송신 신호원은 하나이기 때문에 MUX 를 사용하여 16 개의 송신채널을 하나씩 차례대로 열고 닫게 하였다. 둘째, Rx signal processing unit 에서 12 개의 수신채널을 거쳐 들어온 신호를 하나씩 차례대로 PIC16F877A 에 내장된 A/D 컨버터에 입력하는 것이다. 다음으로 이렇게 입력된 신호의 값을 읽어서 10 비트로 A/D 변환을 하게 되는데 기준전압이 되는 $V_{ref+} = 5V$ 이고 $V_{ref-} = 2.5V$ 이므로 2.5V 일 때 0, 5V 일 때 1023 의 값으로 변환이 된다. 마지막으로 이렇게 A/D 변환된 값을 PC 로 전송해야 하는데 이를 위해서 RS232 통신을 사용하게 된다. 그런데, 실시간으로 맨손의 움직임을 감지하기 위해서는 초당 30 frames 이상을 처리해야 한다. 그렇기 때문에 한 frame 에 해당하는 384 bytes 의 데이터를 초당 30 번 전송해야 하는데 실제로 384 bytes 데이터를 전송하는데 20ms 가 소요되므로 115200bps 의 통신속도를 사용해야 한다. Rx signal processing unit 은 Meshed transparent plate 로부터 입력되는 12 개의 수신 신호를 처리하여 A/D 변환에 적합하게 조정해주는 부분으로 세 가지 기능을 수행하게 된다. 첫째, 수신된 신호는 잡음이 섞여 있는데 우리가 필요한 신호는 400KHz 의 신호이기 때문에 그 외의 다른 신호는 제거해 주어야 한다. 이를 위해서 High Pass Filter 를 사용했다. 둘째, 입력된 신호는 400KHz, 100mVrms 의 Sine-wave 이다. A/D 변환의 기준전압 범위가 2.5V ~ 5V 이므로 현재 입력된 신호를 증폭하고 또한 DC 로 바꾸어주어야 한다. 이를 위해서 OP amp 인 LM318 을 사용하여 증폭 및 정류를 하여 2.5V ~ 5V 이내에서 동작하도록 하였다. 다음은 3D Sensor Table 에 설치된 Signal processing unit 의 prototype 의 사진이다.

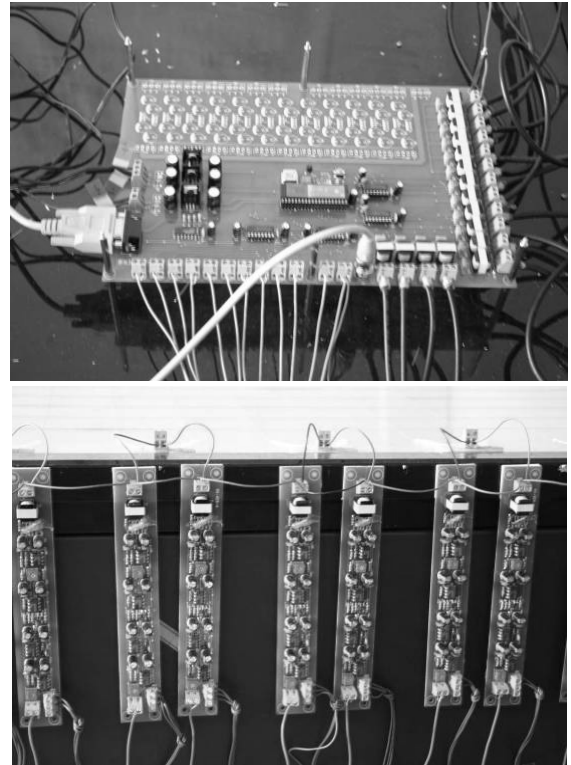


Figure 3-2 Signal processing unit: Master processing unit (up) and Rx signal processing unit (down)

3-1-2 Meshed transparent plate

Meshed transparent plate 는 1140*860*5mm 의 투명한 아크릴판에 6cm 간격으로 세로로 16 개, 가로로 12 개의 폭 1mm, 깊이 1mm 의 홈을 파고 지름 0.3mm 의 copper wire 를 그물망처럼 넣어서 만든 meshed antenna 이다. Signal processing unit 에서 16 개의 송신채널을 통해서 한 채널씩 열고 400KHz, 15Vpp 의 Sine-wave 신호를 흘려주면 해당 송신채널에 대해서 송신 선을 교차하는 12 개의 수신 선이 교차점에서의 손과 meshed transparent plate 와의 proximity 를 측정하게 되는데 이 proximity 값을 통해서 맨손의 위치를 알 수 있게 되는 것이다. 다음은 meshed transparent plate 에서 어떻게 proximity 데이터를 얻게 되는 지 알려주는 그림이다. 그림에서 보듯이 T0 의 송신채널이 On 되고 나머지 송신채널이 Off 일 때 12 개의 수신채널에서 측정되는 수신 값을 차례대로 읽게 되고 이어서 이번에는 T1 의 송신채널만 열고 나머지는 닫아서 T1 에 대한 12 개의 수신 값을 읽게 된다. 이런 식으로 하게 되면 하나의 frame 에 대해서 모두 192 개의 수신 값을 얻게 된다.

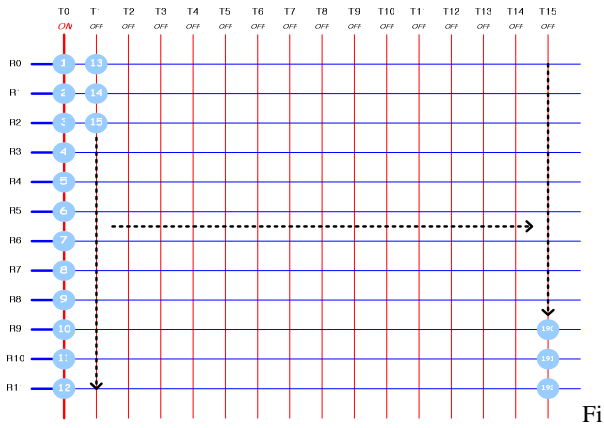


Figure 3-3 Meshed transparent plate sensing process

Meshed transparent plate 는 손과 스크린과의 proximity 값을 측정하기 위한 것인데 기존의 연구에서 사용되었던 방법과는 다른 방식을 사용하여 proximity 를 높일 수 있었다.

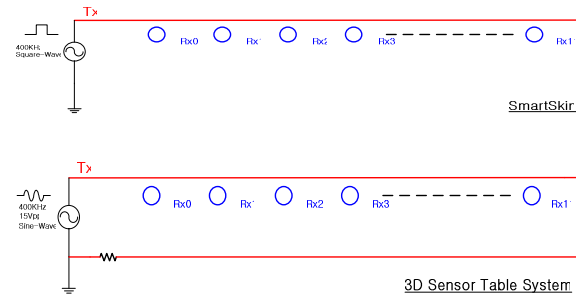


Figure 3-4 Comparison of SmartSkin and 3D Sensor Table System

기존의 연구방법은 Rekimoto 가 2002 년에 제안한 SmartSkin 에서 사용된 mesh-like antenna 인데 그림에서 보는 바와 같이 송신선의 한쪽 끝은 신호원에 물려있지만 다른 끝이 open 되어 있음을 알 수 있다. 그리고 본 연구에서 제안하는 3D Sensor Table System 에서는 송신선이 Loop 를 형성하고 있음을 알 수 있다. 이렇게 Loop 를 사용하였기 때문에 proximity 를 높일 수 있었고 이것을 통해서 스크린에서 30cm 이내의 공간에서 맨손의 움직임을 감지할 수 있는 것이다.

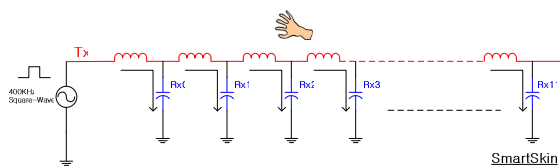


Figure 3-5 equivalent circuit of Tx and Rx of SmartSkin

Figure 3-5 에서는 먼저 SmartSkin 에서 손의 proximity 의 변화에 따른 등가회로를 나타내고 있

다. SmartSkin 은 송신 선중의 하나가 400KHz 의 square-wave 에 의해서 excite 되면 송수신선의 교차점이 마치 capacitor 처럼 동작을 하게 되어 수신선에 신호가 흐르게 되는 것을 이용하였다. 그렇기 때문에 각 교차점에서의 capacitance 에 따라 수신신호의 세기가 변하게 되고 capacitance 는 교차점과 손의 proximity 에 따라서 변하게 된다. 그런데 이 시스템의 문제점은 한쪽 끝이 open 되어 있기 때문에 교차점이 송신 신호원에서 멀수록 그 값이 작아지게 된다. 그래서 proximity 가 5~10cm 정도에서만 측정이 가능하게 되고 3 차원 공간에서 맨손의 움직임을 이용한 application 에서는 사용하기에 적합하지 않다. 반면에 본 연구에서 제안한 방법은 송신선의 한쪽 끝을 open 시키지 않고 loop 로 만드는 것이다. 다음의 그림에서 이 방법에 대한 등가회로를 보여주고 있다.

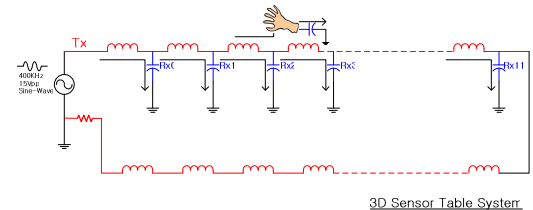


Figure 3-6 equivalent circuit of Tx and Rx of 3D Sensor Table System

이 방법은 송신선이 하나의 loop 를 형성하기 때문에 송신선에 흐르는 400KHz, 15Vpp 의 Sine-wave 를 흘려줄 때, 송신선과 수신선의 임의의 교차점에서 수신선으로 전달되는 에너지가 SmartSkin 과는 달리 안정적으로 일정한 값을 가지게 된다. 그렇기 때문에 스크린으로부터 보다 먼 거리에서도 proximity 를 감지할 수 있게 되었는데 약 30cm 까지 proximity 를 측정할 수 있었다.

3-2 Algorithms

3D Sensor Table 시스템은 meshed transparent plate 로부터 30cm 이내의 공간에서 움직이는 맨손의 위치를 감지할 수 있다. 이것은 다음과 같은 과정을 통해서 수행된다. 먼저, Signal processing unit 에서 rs232 통신을 통해서 20ms 의 속도로 전송하는 384byte 의 데이터를 PC 의 serial port2 를 통해서 받아서 저장한다. 이 데이터는 한 frame 을 구성하

는 192 개의 송/수신선의 교차점에서 얻어지는 raw proximity 이다. 이것을 그대로 사용해서는 제대로 된 proximity 를 알 수 없기 때문에 두 번째 단계에서는 raw proximity 를 proximity calculation 알고리즘을 통해서 modified proximity 를 계산한다. 그러면 이제 proximity 값을 통해서 손의 위치를 찾아야 하는데 그 전에 에러를 제거해야 한다. 왜냐하면, 손의 위치를 찾기 위해서는 proximity 가 최대인 지점을 찾아야 하는데 에러가 발생한 지점이 최대값의 proximity 를 가질 수 있기 때문이다. 에러를 제거한 후에 tracking algorithm 을 사용하여 손의 위치를 구하면 그 다음으로 손의 모양을 찾아야 한다. 현재 만들어진 3D Sensor Table 시스템의 경우에 resolution 이 낮기 때문에 간단한 손 모양만을 인식할 수 있는데 주먹을 쥔 상태와 손을 편 상태를 구분할 수 있다. 마지막으로 이러한 정보들을 바탕으로 3 차원 영상을 만들어서 3D display 에 보내면 3 차원 공간상에서 사람과 컴퓨터와의 인터랙션을 나타낼 수 있게 되는 것이다. 다음의 그림은 signal processing unit 으로부터 넘어온 데이터를 통해서 3 차원 공간에서 손의 위치 정보를 알아내는 과정을 나타내고 있다.

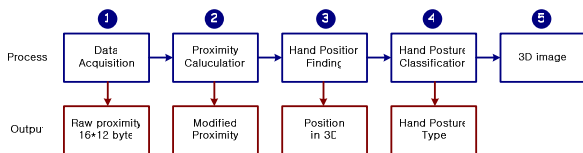


Figure 3-7 the hand-tracking and hand-posture recognition process

Signal processing unit 에서 PC 에 전송된 데이터는 raw proximity data 이다. 3D display 의 스크린 위에서 맨손의 움직임에 따른 192 개의 교차점에서 읽은 proximity 값이지만 그대로 사용할 수가 없다. 이것은 스크린 위에 아무것도 없는 초기상태에서 proximity 값이 192 개의 교차점에서 모두 동일한 값이 아니며 교차점 위에 손이 없을 때 데이터 값이 조금씩 흔들리기 때문이다. 첫째, 교차점 위에 손이 없을 때 데이터 값이 흔들리는 것은 수신되는 신호가 DC 처럼 일정한 것이 아니라 fluctuation 이 일어나기 때문이다. 그래서, Signal processing unit 에서 넘어온 192 개의 proximity data 에 대한

각각의 평균과 최대값, 최소값, 표준편차를 구한 결과 최대값과 최소값의 차이가 10 이하인 것과 표준편차는 5 이하인 것을 확인했다. 따라서, proximity data 를 그대로 사용하는 것이 아니라 평균값을 사용해서 데이터가 흔들리는 것을 줄여줄 수 있었다. 그런데, 데이터 값이 항상 흔들리는 것이 아니라 교차점 위에 손이 없을 때 흔들리며 교차점 위의 30cm 이내의 공간에 손이 있을 때에는 흔들리지 않기 때문에 평균값을 쓸 때 손의 근접으로 인한 데이터의 변동이 문제가 되지 않음을 확인했다.

average	max	min	s.d	range
867	871	863	3.3	8
862	864	858	1.9	6
871	874	866	2.1	6
865	870	860	4.5	10
863	866	858	3.1	8
865	869	860	3.9	9
865	868	860	3.4	8
860	862	857	1.4	5
859	863	856	2.6	7
873	877	868	3.8	9
869	873	865	3.2	8
870	873	866	2.7	7

Table 3-1 example of average, max, min, S.D and range

다음으로 초기상태, 즉, 스크린 위에 아무것도 없을 때의 각각의 교차점에서의 값들이 일정치 않은 문제가 있다고 했는데, 이것은 송신신호와 수신신호가 DC 처럼 일정하지 않기 때문에 발생하는 것으로 보인다. 그래서 12 개의 독립된 Rx signal processing unit 에서 OP amp 의 증폭률을 변화시켜서 A/D converter 에 입력되는 값들의 범위를 일정하게 맞추었고 또, 각각의 교차점에 대해서 상대적으로 초기상태를 맞추고 이후에 들어오는 값들에 대해서 초기값과의 상대적인 차이를 구해서 proximity 를 계산하면 초기상태의 데이터 값이 일정하지 않은 문제점을 해결할 수 있다. 다음의 수식에서 방법을 제시한다.

$$\text{currentM} = \text{previousM} + (\text{previousV} - \text{currentV})$$

Here,

currentV = averaged current input data

previousV = previous value of currentV

currentM = modified value of currentV

previousV = previous value of currentM

Equation 3-1 proximity calculation equation

이렇게 proximity 를 구한 후에 맨손의 위치를 찾게 된다. 맨손의 위치를 알기 위해서는 192 개의

proximity 데이터에서 최대값을 찾아야 한다. 그런데, 때때로 노이즈로 인해 손의 움직임과는 관계 없는 위치에서 최대값이 발생하는 경우가 있다. 그렇기 때문에 이것을 없애기 위해서는 최대값을 갖는 교차점의 근방 교차점들의 값들을 살펴보아야 한다. 실험을 통해서 근방의 값들이 최대값과 5 이하의 차이가 날 때가 예러가 아닌 손의 위치임을 알 수 있었다. 다음의 그림이 이것을 나타내고 있다.

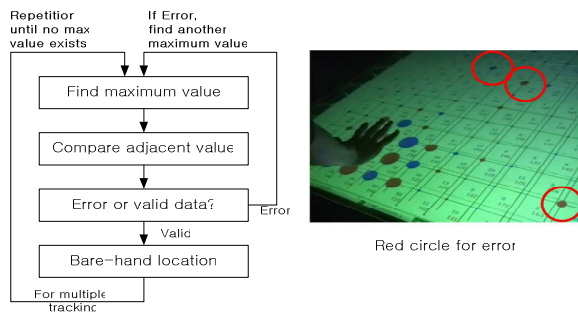


Figure 3-8. tracking algorithm and its implementation

또한, 간단한 손 모양을 인식할 수 있었는데 주먹을 쥔 상태와 손바닥을 편 상태를 구분할 수 있었다. 이것은 손을 폈을 때와 주먹을 쥐었을 때의 손의 위치를 나타내는 교차점의 근방 값들의 분포 모양의 차이를 통해서 알 수 있었다. 다음의 그림에서 이를 보여주고 있다.

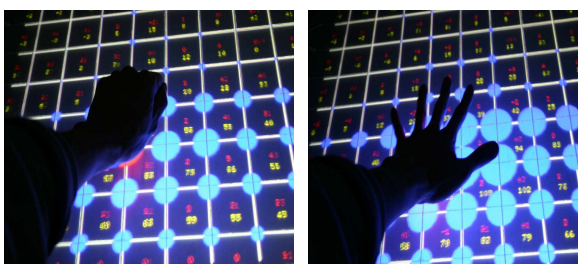


Figure 3-9. When grabbing a hand and stretching a hand

4. 결론

3D 센서테이블 시스템은 30cm 이내의 공간에서 맨손의 움직임을 감지하여 위치정보 및 간단한 제스처를 인식할 수 있었다. Electric Field Sensing 기법을 이용하여 3 차원 공간에서 맨손의 움직임을 감지하는 연구를 찾아보

기 힘든 상황에서 이러한 시스템의 연구는 앞으로 3 차원 애플리케이션을 위한 인터페이스를 개발하는데 도움을 줄 것으로 보인다. 또한, 3D 센서테이블은 2 차원 애플리케이션에서도 사용할 수 있다는 장점을 갖고 있다. 현재로서는 간단한 제스처인식만 가능하지만 트랜시버를 추가하여 정밀도를 좀 더 높이면 다양한 제스처 인식을 할 수 있을 것으로 보인다.

5. 참고문헌

[1] Meredith Ringel, Henry Berg, Yujui Jin, Terry Winograd, Barehands: Implement-Free Interaction with a Wall-Mounted Display, CHI2001, 31 March-5 April, 2001.

[2] Matsushita, N. and Rekimoto, J., HoloWall: Designing a Finger, Hand, Body and Object Sensitive Wall in Proceeding of the ACM UIST'97 symposium on User Interface Software and Technology, 209-210, 1997.

[3] Christian von Hardenberg, Francois Berard. Bare-Hand Human-Computer Interaction, PUI 2001, 2001.

[4] J.R. Smith, T. White, C. Dodge, J. Paradiso, and N. Gershenfeld. Electric Field Sensing for Graphical Interfaces. IEEE Computer Graphics and Applications, 18(3):54-60, 1998.

[5] Jun Rekimoto. SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces, CHI 2002, April 20-25, 2002.

[6] J.R.Smith, Field mice: Extracting hand geometry from electric field measurements, IBM SYSTEMS JOURNAL, VOL35, NOS3&4, 1996.