3D Spatial Touch System
Based on Time-of-Flight Camera

YANG-KEUN AHN, YOUNG-CHOONG PARK, KWANG-SOON CHOI,
WOO-CHOOL PARK, HAE-MOON SEO and KWANG-MO JUNG
Korea Electronics Technology Institute,
#1599, Sangam-dong, Mapo-gu, Seoul
KOREA
ykahn@keti.re.kr

Abstract: - Recently developed Time-of-flight principle based depth-sensing video camera technologies provide precise per-pixel range data in addition to color video. Such cameras will find application in robotics and vision-based human computer interaction scenarios such as games and gesture input systems. Time-of-flight principle range cameras are becoming more and more available. They promise to make the 3D reconstruction of scenes easier, avoiding the practical issues resulting from 3D imaging techniques based on triangulation or disparity estimation. We present a 3D interactive interface system which uses a depth-sensing camera to touch spatial objects and details on its implementation. We speculate on how this technology will enable new 3D interactions. This study implements a virtual touch screen that keeps track of the location of hand, inputted from the disparity image being outputted by a time of flight (TOF) camera, using the Kalman filter. Put out from the depth image of the TOF camera, this image is insensitive to light and therefore helps implement a virtual touch screen independent from the surroundings. The biggest problem with conventional virtual touch screens has been that even the slightest change in location led an object to fall out of or enter the virtual touch screen. In other words, the pointing location responded too sensitively to accurately detect the touch point. The Kalman filter, on the other hand, can be a solution to this problem, as it constantly predicts the pointing location and detects this touch point, without interruption, in response to even a slight change in location. This enables a stable and smooth change in the location of pointing point on the virtual touch screen.

Key-Words: - Depth Sensor, Virtual Touch Screen, 3D Interaction, Kalman Filter, Time-of-flight

1 Introduction

A technology combining display and input device, touch screen is being widely used in various areas. A touch screen, however, is reliant upon a physical membrane and has a size limit; technically speaking, it is vulnerable to vibrations and its maximum optical transmittance stands only at 86%. Against this backdrop, there has been a growing need for the development of technology that can complement these weaknesses of touch screen [1], and virtual touch screen will be the right option for this purpose.

Kim Hyung-joon [1] proposes the use of two cameras for implementing a virtual touch screen. Using two fixed cameras, a virtual touch screen is built based on mathematical locational calculation, not stereo algorithm; when a certain object (i.e. hand) enters the screen, this is recognized as a touch. Martin Tosas and Bai Li [2] implemented a virtual touch screen using one webcam and a physical grid.

When a hand enters the framework of physical grid, as opposed to virtual grid, one webcam tracks the hand, generating an event if a certain motion takes place. Eunjin Koh [3] formed a system structure, using a virtual touch screen—based on the AR marker and two stereo cameras—and HMD; a hand is found and tracked, and then the hand's movements are utilized to implement mouse-like features.

In the technology of Kim Hyung-joon, however, implementation of a touch screen solely based on mathematical locational calculation entails the problem of the screen being fixed mathematically and the touch point being hard to find even when the hand moves only slightly. With only one single webcam being applied, Martin Tosas' technology fails to process three-dimensional depth information; Eunjin Koh's is sensitive to light as stereo cameras are used.

This study, on the other hand, detects the touch point consecutively, using the Kalman filter, despite slight locational change. It also utilizes the TOF camera, where infrared light and infrared camera are used, to extract depth information from an image in a way that is not sensitive to the brightness of light and the surroundings. On this basis, this study suggests a technology for tracking the touch point using the Kalman filter and a technique for configuring a three-dimensional screen, based on depth information, which remains constant against any change in light.

Human computer interaction (HCI) plays a important role in many applications today. It exists the desire to realize more natural forms of interaction between humans and machines. One of these forms is the interaction via gestures, e.g. by recognizing natural hand- or full-body movements. Many researchers have focused on measuring the articulation of the whole body neglecting the hand posture since full-body posture is in itself already a complex task, even when problems due to clothing, self-occlusion and non-rigid tissue can be avoided. On the other hand, the human hand has more that 24 degrees-of-freedom in finger/ thumb articulation and forms a very difficult problem on its own. Solutions to both problems have been developed but they violate an important aspect in HCI, which is called "come as you are". This paradigm is one of the most influential goals in HCI but also the most difficult one to achieve: The intentions of the human user must be recognized effortlessly and non-invasive, e.g. without attaching any special hardware to the human body. Data gloves, suits, inertial tracking devices and optical markers attached to the hand, face or body of the human user must be avoided.

Hence, we present in this paper an approach for non-invasive spatial-touch recognition in real time. We focus on making the hand and its movements useful for human computer interaction.

We decided to investigate the usefulness of the new 3DVsystems Zcam time-of-flight range camera for spatial-touch recognition, as this new hardware exhibits significant advantages over more traditional type of camera sensors: It is able to deliver range data in real time, namely with a frame rate of 30 Hz without the need to solve the correspondence problem, as it is mandatory in depth estimation by means of stereo triangulation. Furthermore no special background or skin color model is required for the segmentation of the hand, so it can be used at

many different places. The camera is based on active IR illumination and is therefore more robust against illumination changes than an optical hand localization based on skin color would be. However, despite these advantages it remains to be investigated in this study how useful the Zcam really is for spatial-touch recognition, where its weaknesses and limits are and how these problems might be investigated by further studies.

2 Depth-sensing Cameras

We refer to camera systems which recover depth information throughout the captured scene (i.e., depth per pixel) as depth-sensing. While we acknowledge the utility of other camera-based means of 3D capture for interactive surfaces, such as recognizing an object from two views, a fully detailed range image permits a great deal of flexibility.

Laser scanners have been used in robotics and other fields to calculate accurate depth images. Despite being available for many years, such technology is still expensive, and often is barely fast enough for interactive applications (Illuminating Clay reports 3Hz scan rate) [19].

The CCD camera in mono or stereo format dominated until recently being perceived as the vision sensor technology of choice. Stereo CCD cameras had a technological advantage edge over depth sensors and have been used as the preferred vision sensor.

A well known problem with CCD stereo vision is the high sensitivity of the depth data to errors in locating corresponding features in each image. Small errors in the contrast areas limit in the two image results in significant depth measurement error.

Correlation-based stereo is another old approach which suffers from a number of difficulties. For example, stereo matching typically fails on regions of the scene with little or no texture. Secondly, even today stereo matching requires a great deal of computational power to obtain interactive rates at reasonable resolution. Finally, stereo camera setups typically require fine calibration.

The newly developed CMOS 3D sensors provide depth data in addition to reflected light intensity. They have become an important contender as the most frequently used sensor type.

Recent advances in CMOS Time of Flight (TOF) 3D infrared sensor increased interest in their use in

real time visual perception applications. Fabricated using the standard CMOS process integrated circuits 3D sensors are now available commercially at a very low cost.

The 3D TOF sensors have numerous advantages over other depth sensors. Triangulation-based methods such as stereo cameras require intensive post processing to construct depth images. This is not necessary in the case of the TOF sensor, and the post processing usually involves a simple table lookup to map the individual pixel sensor reading to real range data.

The operation of the sensor is based on an amplitude-modulated infrared light source and an array of CMOS transistors that determine the field depth from the back scattered light. The ambient light is not affecting the sensor operation since it is not modulated. Several other methods are used to eliminate the effect of ambient noise.

The camera module contains a light source constructed from a bank of infrared LEDs at 870nm wavelengths, a lens system for the detector chip incorporating 320×240 phase-sensitive pixels.

The whole chip is fabricated on standard CMOS process and the chip also includes an embedded processing unit for pixel depth value calculation.

The time-of-flight (TOF) sensor measures distances by determining the time delay between emission and detection of light pulses. The pulses are emitted by the light source switched on and off with a 50% duty cycle at a frequency on the order of 50 MHz. A simple calculus shows that the distance travelled by light in an impulse period is about 3 meters. At a digital signal resolution of 8 bits the depth resolution is on the order of one centimeter.

The frequency of the light source defines the maximal depth of the field of view. There are methods using multiple beams of different frequencies to extend the sensor range up to 100 m.

The light beam bounces off surfaces in the scene and returns to the 3D sensor. The phase of the reflected beam and geometrical position of light source to each pixel are used to determine the depth information in the 3D scene viewed.

The core of the sensor design is a pixel matrix consisting of special CMOS transistors with two gates. The differential structure accumulates photon generated charges in the two collecting gates. The gate modulation signals are synchronized in quadrature with the light source, and hence depending on the phase of incoming reflected light, one node collects more charges than the other. An

integration in time over many cycles is performed to increase sensitivity. At the end of the integration cycle the voltage difference between the two nodes is read out as a measure of the phase of the reflected light.

New camera technologies under development observe depth information in a more direct fashion, and so address many of the drawbacks of previous approaches [18][20].

In the present work we use the Zcam camera by 3DV Systems, Ltd [9]. The Zcam times the return of pulsed infrared light: reflected light from nearer objects will arrive sooner. A Gallium-Arsenide (GaAs) solid-state shutter makes this possible (see Fig. 1). The result is an 8 bit depth image, over a variable dynamic range (70cm to 3m). Fig. 1d illustrates an example Zcam depth image. The Zcam camera also includes a separate color camera. The output of the color camera is registered with the depth image to obtain a complete "RGBZ" image at 30Hz.
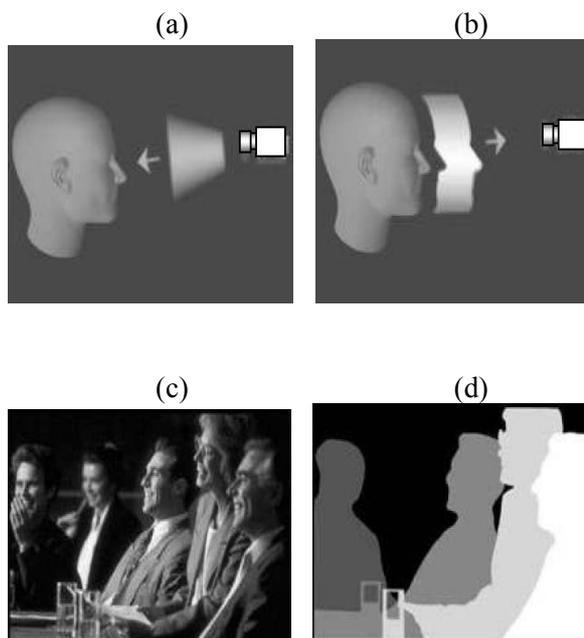
(a)   (b)

(c)   (d)

Fig. 1. 3DV Zcam uses pulsed infrared light (a) and solid state shutter to compute depth image. Reflected light from closer surfaces arrives sooner (b). Example real view (c) and depth image (d).

2.1 The Time-of-Flight Principle

A time of flight (TOF) camera calculates the distance from an object by sending out laser or infrared LED and determining—with its built-in sensor—the time spent for laser/infrared LED particles to hit the object and come back [6]. If the laser or infrared LED that the camera emits forms a sine curve, the equation of the sine curve, s(t), can be written as Equation (5) when modulated frequency is f_m.
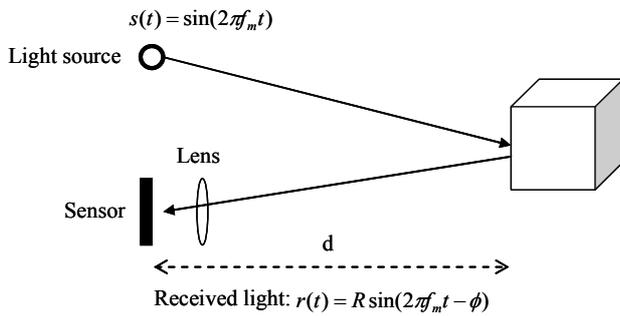

Fig. 2 how a TOF camera works [7]

$$s(t) = \sin(2\pi f_m t) \qquad (5)$$

Then, r(t), which represents the light reflected by the camera and the object and transmitted into the built-in sensor, can be expressed as Equation (6).

$$r(t) = R\sin(2\pi f_m t - \phi) = R\sin(2\pi f_m(t - \frac{2d}{c}))$$
$$(6)$$

Here R is the magnitude of the reflected light and ϕ is the value of the light phase-changed from the object; c is the speed of light, that is, $3 \times 10^8 \, m/s$; and d is the distance between the object and the camera, which can be obtained from the phase-changed value, and this can be written as Equation (7).

$$d = \frac{c\phi}{4\pi f_m} \qquad (7)$$

Using the value of distance obtained above, the TOF camera determines the disparity value (i.e. distance). Fig. 3 illustrates a disparity map
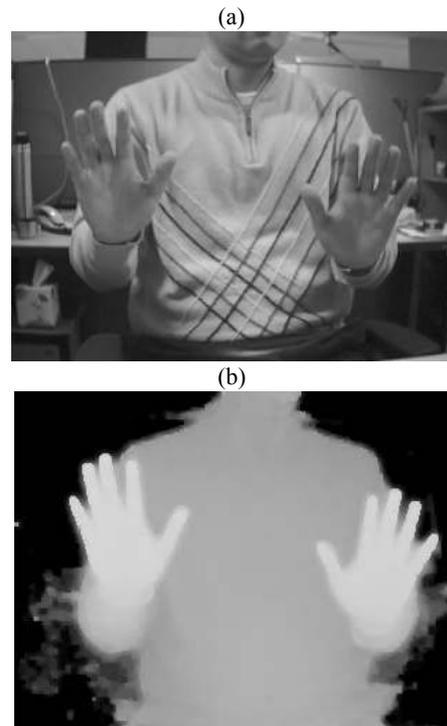
generated after identifying the distance.

(a)

(b)

Fig. 3 (a) input image; (b) disparity image

2.2 TOF Camera

Time-of-flight principle 3D cameras are becoming more and more available and their acquisition costs are continually decreasing. The experiments described herein were conducted using a depth-sensing camera called Zcam built by the 3DVsystems.

Time-of-flight principle cameras work similar to radars. The camera consists of an amplitude-modulated infrared light source and a sensor field that measures the intensity of backscattered infrared light. The infrared source is constantly emitting light that varies sinusoidal between a maximum and a minimum. Objects of different distances to the camera reflect the light in different intensities. The reason is that, at the same moment, objects that have different camera distances are reached by different parts of the sinus wave.

The incoming reflected light is then compared to the sinusoidal reference signal which triggers the outgoing infrared light. The phase shift of the outgoing versus the incoming sinus wave is then proportional to the time of flight of the light reflected by a distant object. This means, by measuring the intensity of the incoming light, the phase-shift can be calculated and the cameras are

able to determine the distance of a remote object that reflects infrared light. The output of the cameras consists of depth images and a conventional low-resolution gray scale video, as a byproduct. A detailed description of the time-of-flight principle can be found in [15, 16, 17].

The depth resolution depends on the modulation frequency. For our experiments we used a frequency of 20MHz which gives a depth range between 0.5m and 7.5 m, with a theoretical accuracy of about 1 cm. Usually, time-of-flight cameras allow to configure frame rate, integration time, and a user-defined region of interest (ROI) by writing to internal registers of the camera. The cameras then calculate the distances in an internal processor. The resolution of the Zcam camera is 320×240 non-square pixels. Unfortunately, 3D time-of-flight cameras are not yet available in higher resolutions such as NTSC or PAL.

2.3 Zcam System Operation

The system captures the depth and RGB values of each pixel in the scene, thus creating a depth map of the object by gray level scaling the distances. And so information such as the object boundaries; object surface elevations and object distances is provided as an output, at standard video rate.

As the technology measures depth linearly, the depth accuracy generated by the system is constant regardless of the distance from the camera - one could slice the measured space and measure each slice at a high depth resolution. The combination of the results yields a high-resolution depth representation of the entire scene.
For applications that use the depth information in order to construct a 3D representation of the scene, such as virtual lighting of the scene, a simplified depth map can be used to reduce the data bit rate and the amount of calculations needed to generate a surface in 3D animation programs.
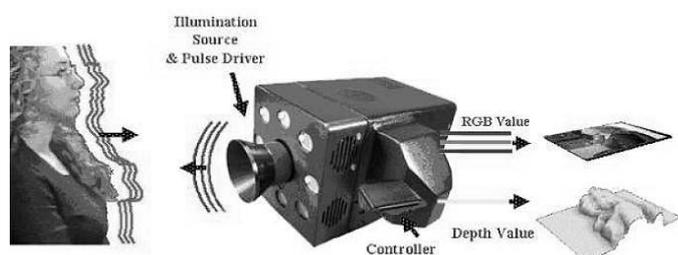

Fig. 4 Zcam System Operation

3 Overview of Kalman filter

Filtering has continuous nature: To estimate the state at the time of t requires the data up to t-1; once the data at t are obtained, the state at the time of t+1 is estimated on this basis. One sort of filtering technique, the Kalman filter has the characteristics of continuous optimality that seeks for the most optimal forecast at each time of estimation; forming a model in the time area, it is easy to deal with mathematically and is therefore being widely used.

The basic premise of Kalman filter is that observation values do have errors. The goal here is to process the observation values—with errors included—using the Kalman filter and renew the estimated values to make an optimal estimation.

The Kalman filter can be broadly divided into three parts: system equation, measurement equation and Kalman filtering. In the system equation part, state vector X(t), which indicates dynamic properties by time, to estimate the state of the system.

$$X(t) = \phi(t-1)X(t-1) + w(t-1) \quad (1)$$

In Equation (1), X(t) is the state vector at the time of t and X(t-1) the state vector at the time of t-1. ϕ(t-1) is a transition matrix that transits the state vector at t-1 into one at t. w(t-1) is a system error vector that obeys normal distribution with the mean of 0 and variance of Q.

As shown in Equation (2), the measurement equation represents the relation for deriving observation vector Z(t) from state vector X(t).

$$Z(t) = H(t)X(t) + v(t) \quad (2)$$

In Equation (2), Z(t) is an observation vector and H(t) an observation transition function matrix that transits state vector X(t) into the observation vector. v(t) is an observation error vector which follows normal distribution with the mean and variance of 0 and R, respectively.

System equation in the Kalman filter is the process of renewing the initially estimated values at the time of t using observation values.

$$\hat{x}(t \mid t) = K'(t)\hat{x}(t \mid t-1) + K(t)z(t \mid t-1) \quad (3)$$

In Equation (3), $\hat{x}(t \mid t)$ is an optimal estimated

value at the time of t and $\hat{x}(t \mid t-1)$ is a state estimation value at the time of t-1, which is estimated using the data up to t-1. K'(t) and K(t) are time-varying weighted matrices enabling state estimation values at each point of time to have minimum variance. When K(t) is calculated and selected as in the below Equation (4):

$$K(t) = P(t \mid t-1)H^T(t)\left[H(t)P(t \mid t-1)H^T(t) + R(t)\right]^{-1} \quad (4)$$

Optimal estimation $\hat{x}(t \mid t)$ can be obtained, where P(t|t-1) is error covariance. Fig. 5 visualizes this process of renewal [4].
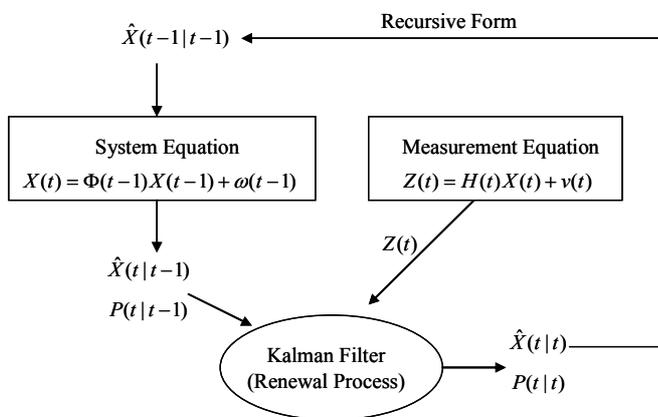
Fig. 5 renewal process of kalman filter [5]

4 Implementation of Virtual Touch Screen

This study implements a virtual touch screen, tracking the touch point with the Kalman filter and using the TOF camera.

4.1 Determination of threshold value

The very first step to implement a virtual touch screen is to separate the hand's location from the inputted depth image of TOF camera by setting the threshold value on the image. Fig. 7 shows depth and x-depth images. In an x-depth image, the value of the x-axis on the image refers to the x-axis as in a real image and its y-axis stands for the biggest among the depth values of the y value. In other words, this represents the silhouette seen by the camera from above.

In Fig. 7, (a) is the depth image when the actor does not keep his hand or hands forward, and (b) is the x-

depth image for the same situation. What sticks out here represents the whole of the actor's body. Presented in (c) is the depth image of a hand brought forward. In comparison to (b), the hand, when seen from (d), sticks out more than the whole of the actor's body. By applying a certain threshold value to this x-depth information, x-depth can be separated into hand, body and background. Fig. 7 visualizes this hand-body-background separation based on the x-depth image. In this figure, (a) and (c) are depth images; the upper part of (b) represents the body of the actor and the lower part the background. Shown in (d) are the results of detecting the hand areas other than the body and background of the actor by setting the threshold value.
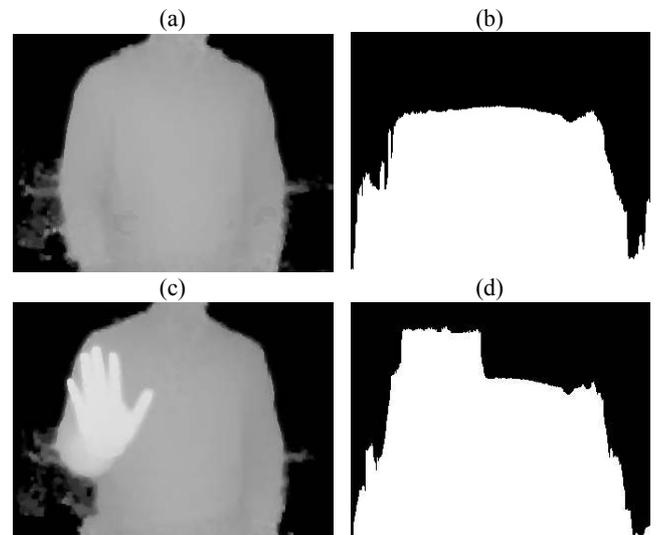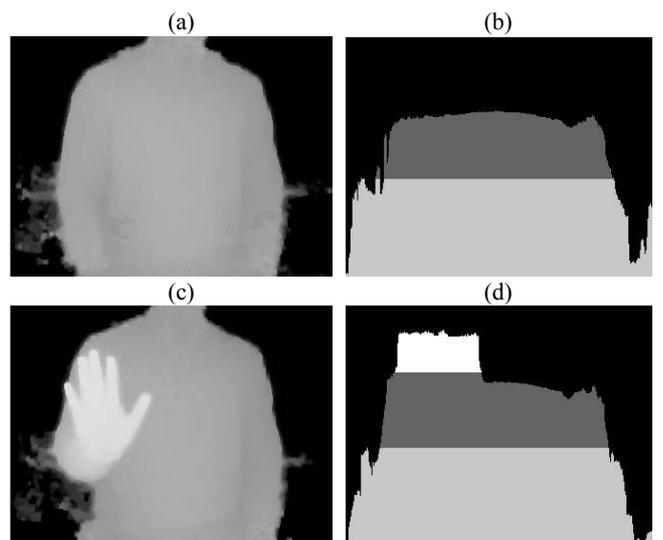

Fig. 6 depth and x-depth transformation images

Fig. 7 x-depth image where threshold value is applied

4.2 Labeling

Once the hand areas are detected using the threshold value, what is needed now is to put together meaningful parts of those areas (i.e. labeling process). If only one hand is completely detected, this process would of course be unnecessary; when both hands are detected, labeling should be done to recognize each of them separately or eliminate the errors of these two hands joined together. For the labeled hand areas, the central moment value should be used to find out the central point again, which will then be used in updating observation vector Z(t). Fig. 8 shows the central point of the extracted hand areas—expressed in the form of a circle—that are labeled and whose central moment values are calculated.

Fig. 8 central point after calculation of central moment

4.3 Tracking of touch point

After labeling and calculation of central moment, the Kalman filter is used to apply the central point obtained this way to observation vector Z(t) and thereby track the touch point.
State vector X(t) is set using the x- and y-axes and

their respective amounts of change in the image labeled from the camera; observation vector Z(t) is determined using the values of the x- and y-axes observed from the image inputted real-time. Also, transition matrices (t-1) and H(t) are set as described in Equation (8).

$$\phi(t-1) = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, H(t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

(8)

Shown in Fig. 9 is the image of touch points being tracked, using the Kalman filter where state and observation vectors are applied. The large circle represents the central point of a real hand, and the small circle is one measured by calculating the noise of the Kalman filter into the hand. The rectangular is the point forecast by the Kalman filter.
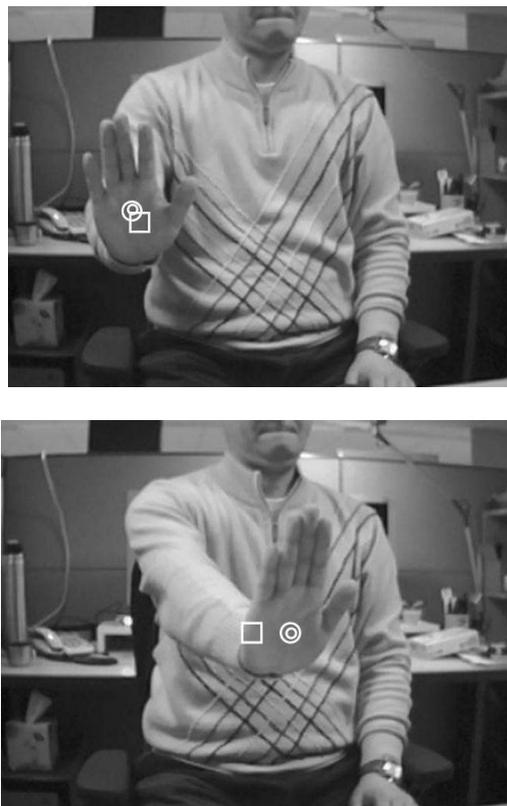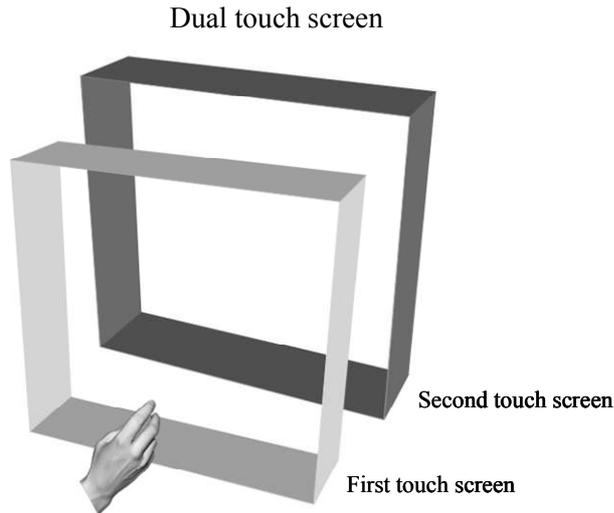

Fig. 9 tracking based on Kalman filter

Dual touch screen

Second touch screen

First touch screen

A virtual touch screen is implemented using two threshold values, so it has the structure of dual touch screen. The primary virtual touch screen keeps track of the touched x and y values and corresponds them to mouse movement events so that the movements of the mouse can be controlled on the screen. The secondary screen is devised to recognize a touch, which sets off an event equivalent to a mouse click. Fig. 10 illustrates the input structure of a dual-structured virtual touch screen.

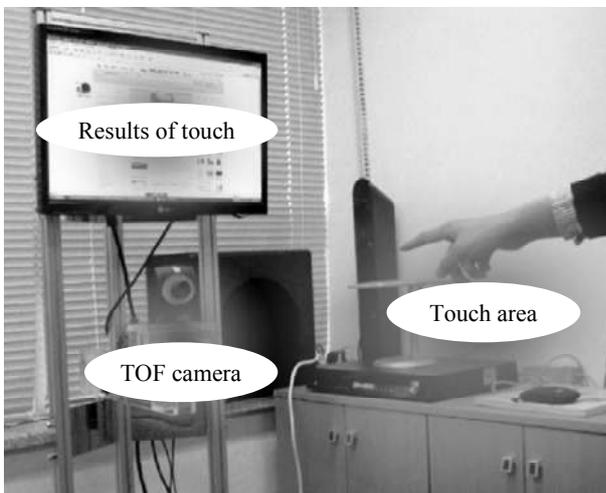Results of touch

Touch area

TOF camera

Fig. 11 actually implemented virtual touch screen

Fig. 11 is the picture of an actually implemented virtual touch screen, which has the actor area that receives gesture inputs from the actor; the camera area receiving inputs of what the actor does; and the screen part that runs the touch recognition process and shows the results of touch recognition.

5 Implementation of 3D Interaction

5.1. System Configuration

To explore the application of depth-sensing cameras to interactive 3D interface, we built a spatial-touch system which uses the camera to recover center position of human hand. Virtual objects to manipulate are projected on screen. We are interested in the range of interactions made possible by this configuration. To begin, we implemented a simple 3D Spatial-Touch interface using Holomatix Blaze 3D Studio (Real-time Interactive Rendering Ray-tracing Software). Users manipulate virtual objects displayed on the screen. The user can rotate virtual objects by hand. Holomatix 3D Studio is used to display 3D virtual objects, for example, the 3D cellular phone rendering model.
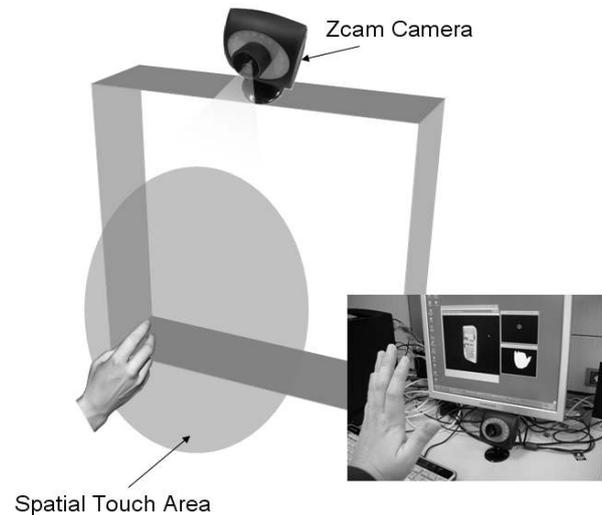

Zcam Camera

Spatial Touch Area

Fig. 12 System Configuration

5.2 3D Interaction System

The 320x240 depth map image is returned by the camera at a frame rate of 30Hz. The camera is configured to place its depth-sensing dynamic range at front of the display.

Because the focal length of the camera is known, it is possible to calculate the 3D position (in centimeters, for example) indicated at each pixel in depth image. It is straightforward to then construct a vertex buffer for this depth map, and it is similarly easy to texture map this mesh with the color image also returned by the camera.
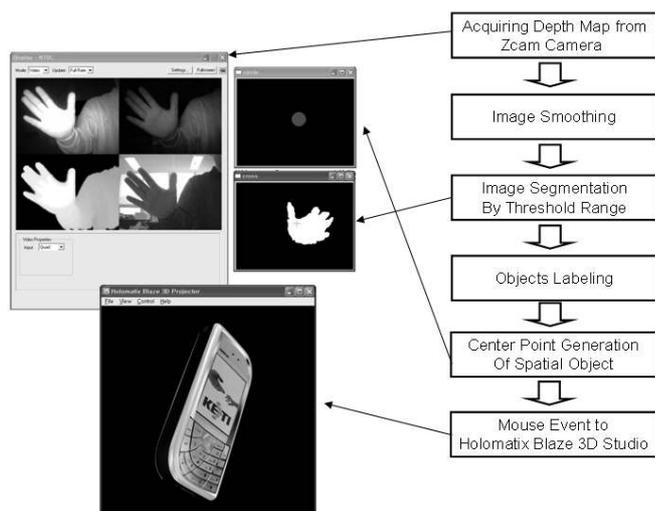
Fig. 13 3D Interaction

Fig. 14 3D Interaction with Virtual Object

The Zcam depth image is somewhat noisy. For some applications it will be necessary to smooth the image to mitigate the effects of shot noise. We are able to reduce speckle noise significantly by applying a median filter to the range image.

Arm and hand are segregated from the background by simple depth keying, e.g. we define a region of interest (ROI) in depth and discard any measurements which do not fall within the predefined range of depth.

The depth range defines also a working point for our recognition system as it marks the optimal trade off between two conflicting properties: The hand should cover as much area of the camera sensor as possible, but any overexposure must be avoided. Overexposure results in serious errors as it impairs the depth measurements of the complete scene significantly.

We use a fast labeling technique (GrassFire labeling) which allows an easy human hand's center point generation process. Fig. 13 and 14 show 3D

spatial-touch software flow and our application.
It is interesting to watch people attempt to interact with the virtual objects directly with their hands.

6 Conclusion

We have described a research of doing 3D spatial touch recognition with a novel IR time-of-flight range camera. After image segmentation and object labeling to get useful measured data we defined a region of interest and interpreted the remaining data as a point cloud. The system was able to recognize spatial-touch of human hand with 2-3 Hz frame-rate. This is a promising result and defines a road map for further research in the future.

This study has built a virtual touch screen system, using the touch point tracking technique based on the TOF camera and the Kalman filter. To address the widespread problem in conventional virtual touch screen systems (i.e. responding sensitively to even the slightest hand movement and failing to recognize the touch point accurately), this study has applied the Kalman filter tracking technique to predict and track the touch point consecutively. Being forecast and recognized, the predicted touch point is bound to exist all the time, but an abrupt movement of touch point is likely to be recognized only insensitively. In future studies, an algorithm that reacts sensitively to a sudden touch point movement needs to be applied to develop a virtual touch screen responding more spontaneously to gestures.

Acknowledgements

References:
[1] Kim Hyung-joon, "Virtual Touch Screen System for Game Applications", Journal of Korea Game Society, vol. 6, no. 3, pp77-86, 2006.
[2] Martin Tosas and Bai Li, Lecture Notes in Computer Science, Heidelberg, Springer Berlin, pp48-59, 2004.
[3] Eunjin Koh, Jongho Won, and Changseok Bae,"Vision-based Virtual Touch Screen Interface", Proceeding of ICCE 2008, LasVegas, USA, 2008.

[4] Chang Seok-kon, "Combining Forecasts of Chunf-Ju Dam Daily Inflow Using Kalman Filter", 2002.

[5] Eric Cuevas, Daniel Zaldivar, and Raul Rojas, "Kalman Filter for vision tracking", 2005.

[6] http://en.wikipedia.org/wiki/Time-of-flight

[7] Gokturk. S. B, Yalcin. H, and Bamji. C., "A Time-Of-Flight Depth Sensor - System Description, Issues and Solutions", CVPRW '04, p35, 2004.

[8] Kalman, R. E. (1960), A New Approach to Linear Filtering and Prediction Problems, Journal of Basic Engineering, Vol. 82, No. 1, pp. 35-45.

[9] G. J. Iddan and G. Yahav, "3D Imaging in the Studio," SPIE, vol. 4298, pp. 48, 2001.

[10] B. Piper, C. Ratti, and H. Ishii, "Illuminating Clay: A 3-D Tangible Interface for Landscape Analysis," presented at CHI 2002 Conference on Human Factors in Computing Systems, 2002.

[11] T. Starner, B. Leibe, D. Minnen, T. Westeyn, A. Hurst, and J. Weeks, "The perceptive workbench: Computer-vision-based gesture tracking, object tracking, and 3D reconstruction for augmented desks," Machine Vision and Applications, vol. 14, pp. 51-71, 2003.

[12] A. Wilson, "TouchLight: An Imaging Touch Screen and Display for Gesture-Based Interaction," presented at International Conference on Multimodal Interfaces, 2004.

[13] A. Wilson, "PlayAnywhere: A Compact Tabletop Computer Vision System," presented at Symposium on User Interface Software and Technology (UIST), 2005.

[14] A. Wilson, "Robust computer vision-based detection of pinching for one and two-handed gesture input," presented at UIST, 2006.

[15] S. B. G̈okẗurk, hakan Yalcin, and C. Bamji. A Time-Of-Flight Depth Sensor - System Description, Issues and Solutions. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Washington D.C., USA, July 2004.

[16] X. Luan, R. Schwarte, Z. Zhang, Z. Xu, H.-G. Heinol, B. Buxbaum, T. Ringbeck, and H. Hess. Three-dimensional intelligent sensing based on the PMD technology. Sensors, Systems, and Next-Generation Satellites V. Proceedings of the SPIE., 4540:482?487, December 2001.

[17] T. Oggier, M. Lehmann, R. Kaufmann, M. Schweizer, M. Richter, P. Metzler, G. Lang, F. Lustenberger, and N. Blanc. An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger). Optical Design and Engineering. Proceedings of the SPIE., 5249:534–545, Februar 2004.

[18] I. Jivet and A. Brindusescu, Real Time Representation of 3D Sensor Depth Images, WSEAS Transactions on Electronics, Issue 3, Vol 5, March 2008 pp:65-71

[19] M Otesteanu, V Gui, 3D Image Sensors, an Overview, WSEAS Transactions on Electronics, Issue 3, Vol 5, March 2008 pp:53-56

[20] K Boehnke Fast Object Localization with Real Time 3D Laser Range Sensor Simulation, WSEAS Transactions on Electronics, Issue 3, Vol 5, March 2008 pp:83-92