International Journal of Engineering Research
Volume No.3 Issue No: Special 2, pp: 103-106
ISSN:2319-6890)(online),2347-5013(print)
22 March 2014

Refined 3d Hand Tracking and Detection Using Kalman Filter

Palli Padmini, M.Tech (DECE), KSRM College of Engineering, Kadapa

Abstract : Hand gestures are an important type of natural language used in many research areas such as human-computer interaction and computer vision. Hand gestures recognition requires the prior determination of the hand position through detection and tracking. One of the most efficient strategies for hand tracking is to use 2D visual information such as color and shape. However, visual-sensor-based hand tracking methods are very sensitive when tracking is performed under variable light conditions. Also, as hand movements are made in 3D space, the recognition performance of hand gestures using 2D information is inherently limited. In this article, we propose a novel real-time 3D hand tracking method in depth space using a 3D depth sensor and employing Kalman filter. We detect hand candidates using motion clusters and predefined wave motion, and track hand locations using Kalman filter. To verify the effectiveness of the proposed method, we compare the performance of the proposed method with the visual-based method. Experimental results show that the performance of the proposed method out performs visual-based method.

Keywords: HCI, Kalman, *information*

1. Introduction

Recently, human-computer interaction (HCI) technology has drawn attention as a promising man-machine communication method. Advancements of HCI have been led by associated developments of computing power, various sensors, and display techniques. Interest in human-to-human communication modalities for HCI also has been increased. These include movements of human hands and arms. Human hand gestures are non-verbal communication that ranges from simple pointing to complex interactions between people. Main advantage of hand gestures is the ability of communication in the distance. The use of hand gestures for HCI demands that the configurations of the human hand can be measurable by the computer. The performance highly depends on the accuracy of detection and tracking of hand locations.

Current hand detection and tracking methods are using various sensors including directly attached to hand, special feature gloves, and color or depth images.The hand detection and tracking via image sensor may be done with 2D or 3D information. However, as obtaining 3D information needs high computing power and high cost equipment, 2D methods have been more developed than 3D. In 2D hand detection and tracking methods, the most common method is a visual-based method, which uses information such as color, shape, and edge. Visual-based methods can be categorized as color-based and template-based methods. The color-based method starts by finding a hand region using color information (RGB, HSV, YCbCr). Then, a color histogram is made from the detected hand. Based on this color histogram the region which is similar to hand color can be tracked .The template-based method creates an edge image through the color or gray image. The edge image is matched to the trained hand template, and then the hand is tracked. However, hand movements generally occur in 3D space. Then, 2D method only can use 2D information, which eliminates the movement information along the z-axis. This makes the limitation of 2D methods inherently. Recently, the equipment for obtaining 3D information is becoming faster, more accurate, and cost-effective. This equipment includes depth sensors such as TOF cameras and Prime Sensor. After the emergence of this equipment, real-time 3D hand tracking methods rapidly developed. For example, Breuer et al. used an infra-red TOF camera to create a near real-time gesture recognition system. Grest et al proposed a human motion tracking method using a combination of depth and silhouette information. In this article, we propose a novel real-time 3D hand tracking method in depth space using Prime Sensor with Kalman filter.

We generate the motion image from depth image. Then, we detect hand candidates using motion clusters and predefined wave motion, and track hand locations using Kalman filter. The organization of this article is as follows. related works are briefly reviewed. The preprocessing of depth information and the proposed hand detection and tracking method are described. Several experiments of our hand tracking system are performed.

2. Background

2.1 Visual-based hand tracking

There are two well-known visual hand tracking methods: color- and template-based methods. In color-based methods, after initial hand detection, the color information is extracted from the specified initial region. This color information is made up of RGB-space pixel colors or transformed into HSI-space pixel colors. In the color histogram is made from hue and saturation values of the region. Then, the obtained color histogram is used to hand tracking. In template-based methods, the initial hand is found by matching the whole image with a prepared trained hand template. The template is moved near to the initial hand region, and the matching point of the hand is found. This process is used for every frame. Visual-based methods are natural tracking method. However, visual-based methods are highly affected by the illumination conditions. When using a color histogram or skin color probability density function, RGB, hue, and saturation values may change by illumination. This can make it difficult to find and track the hand. Also, when a specific part of the hand is occluded or shaded by an object, then hand tracking can fail.

2.2. Depth-based hand tracking

Depth-based hand tracking methods can be categorized into model-based and motion-based. Model-based hand tracking uses the 3D articulation model to fit the hand. The motion-based method uses hand motion in depth space.

Breuer et al proposed the model-based hand tracking in depth space. In order to estimate location and orientation of the hand, principal component analysis is used with 3D points. They optimized hand model parameters to minimize discrepancy between the appearance and 3D structure of hypothesized instances of a hand model and the actual hand observations. The tracker based on stochastic meta-descent for optimizations in high dimensional state spaces is proposed by Bray et al. This algorithm is based on a gradient descent approach with adaptive and parameter-specific step sizes. The hand tracker is reinforced by the integration of a deformable hand model based on linear blend skinning and anthropometrical measurements. In motion-based hand tracking method, Holte et al. proposed the view invariant gesture recognition system with the TOF camera. This method finds the motion primitives from an accumulated image based on 3D data. It detects movements using a 3D vision of 2D double differencing (subtracting the depth values pixel wise in two pairs of depth images), thresholding, and accumulating.

2.3. Color information versus depth information

Figure 1 shows the color and depth images under different illumination conditions. Figure 1a,b shows the color and depth images with normal illumination condition. In contrast, Figure 1c,d shows them in low illumination condition.

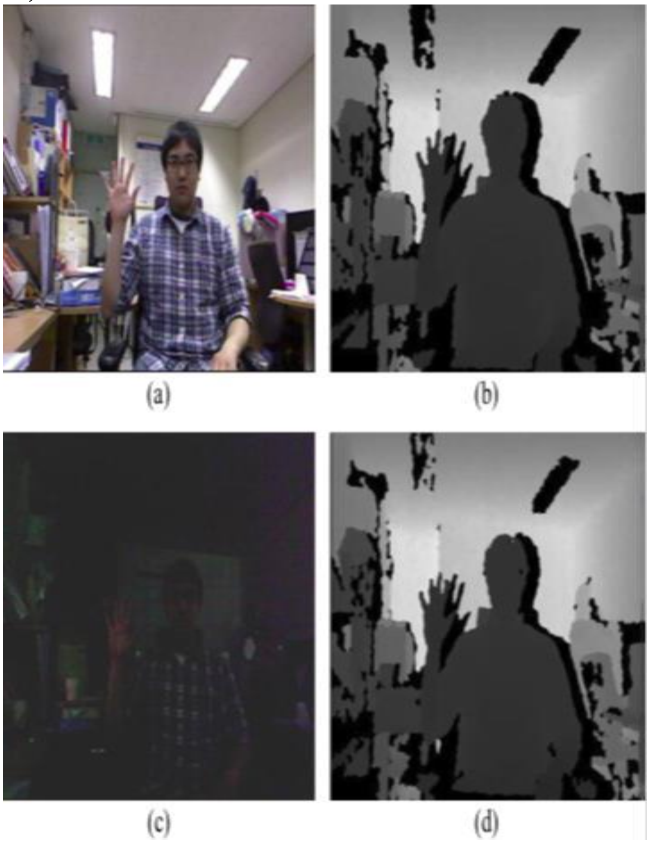

Figure 1 Comparing color and depth images under different illumination conditions. (a) color image in normal illumination; (b) depth image in normal illumination; (c) color image in low illumination; (d) depth image in low illumination

The figures show the sensitivity to illumination changes of color and depth images. As figures showing, the color image is very sensitive to illumination variation. The TOF camera and the Prime Sensor are currently developed depth image sensors. Both sensors produce depth images that store the real depth value in each pixel. For example, the Prime Sensor stores in each pixel with 16 bits depth information. We have the image with 3D information X, Y, and Z-axis. The depth image also has some drawbacks. First, the depth image includes a lot of noise at the edge of objects. Second, it is hard to find invariant features of objects, because the depth information depends only on distance.

2.4. Kalman filter

Kalman proposed a recursive method to solve the problem of linear filtering of discrete data. Providing many advantages in digital computing, Kalman filter is applied in a variety of research fields and real application areas.The main procedure of Kalman filter is to estimate the state, then refine the state from the
error.

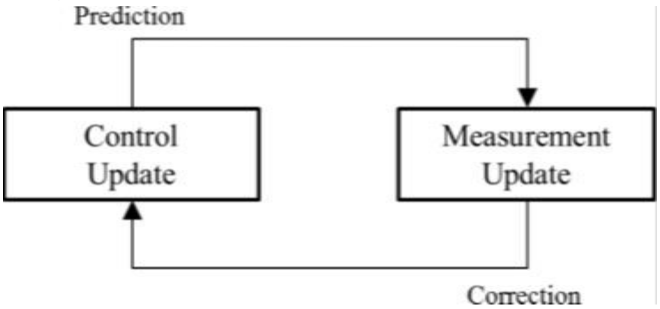

Figure 2 The procedure of Kalman filter.

The Kalman filter has two update procedures as shown in Figure 2. One is a control update and the other is a measurement update. In the control update, we estimate the state with the previous state and an action parameter (vector). In the measurement update, the state is corrected by sensor information.

3. Proposed method

In this section, we explain the proposed hand detection and tracking algorithm. Figure 3 shows the steps of the proposed method. First, we get a depth image from the depth sensor, and create a motion image which is the accumulated difference images. Then, we reduce the noise with the spatial filter and the morphological
operation. Motion clustering method is proposed to find motion clusters. Then, initial hand detection is performed among the clusters with wave motion. Finally, the Kalman filter is used to track the hand.

3.1 Motion image

We use the motion image which is the accumulated difference image. We store consecutive images in the chronological order. Then, we obtain the difference image which is the previous frame (it-1) subtracted from the current frame (it).

We accumulate difference images .In this accumulated image, all movement of human, object, and noise are represented. Next, noise reduction, motion clustering, and hand detection procedures are applied to this motion image.

3.2 Noise reduction

We use a spatial filtering and a morphological processing for noise reduction. This median filter provides excellent salt and pepper noise reduction with considerably less blurring.

International Journal of Engineering Research ISSN:2319-6890)(online),2347-5013(print)
Volume No.3 Issue No: Special 2, pp: 103-106 22 March 2014

We also use morphological processing for noise reduction. The erosion operation slips off the object or particles layer, reducing irrelevant pixels and small particles from the image. The dilation operation does the inverse of the erosion operation.

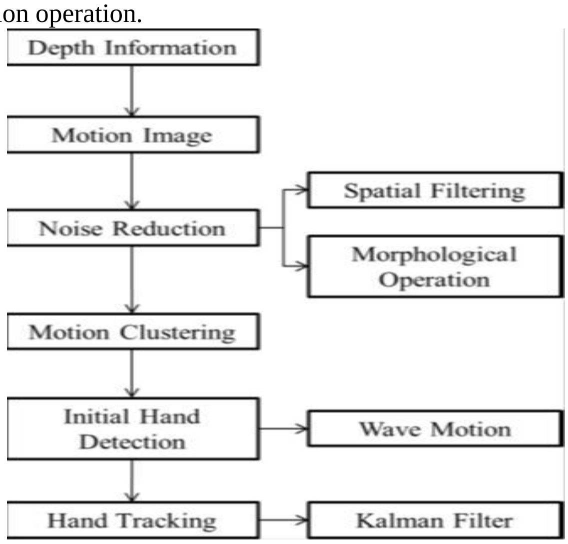

Figure 3 Procedure of the proposed hand tracking method

3.3 Motion clustering

First we select connected components from the motion image. Then the obtained connected components are clustered. These clusters are possible candidates for the hand. The selected clusters can be either real motion or noise. The noise clusters are usually small or split frequently, so if the size is smaller than some threshold, then we can decide it as a noise cluster, and remove it.

3.4 Initial hand detection

To find the hand, we set the condition of hand wave motion, which consists of a side-to-side motion sequence. First, we detect the direction of cluster movements using a motion template. The motion template is an effective method for tracking general movement, and it is especially useful for gesture recognition.

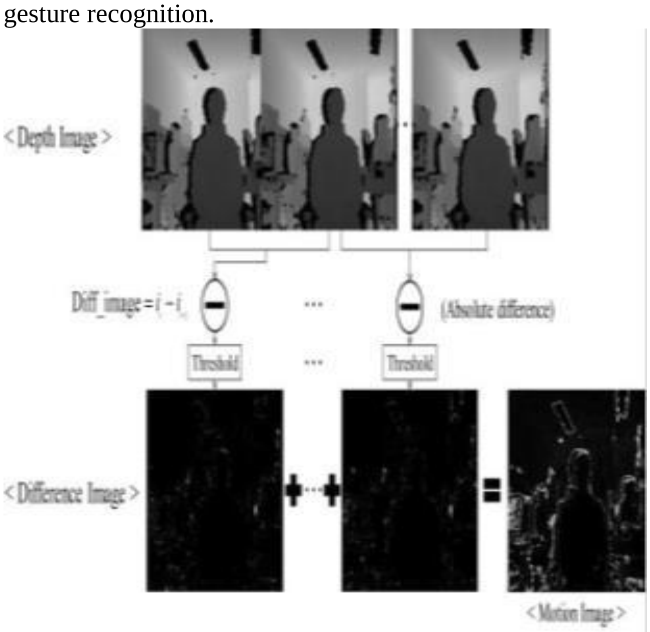

Figure 4 The process of generating the motion image

3.5 Hand Tracking

The Kalman filter needs small data storage for previous data in operating the recursive process, because we only need information of the previous state, and not the whole previous frame.

The second advantage is that the Kalman filter is suitable for treating a time varying signal. Therefore, we apply the Kalman filter for hand tracking. The Kalman filter is used for object tracking in many applications. Usually they use the state which is two dimensions for visual images. But as we need to add depth information, the state is designed as three dimensions.

We assign depth information as z-axis values. Using the depth axis in Kalman filter tracking, we can track the hand more accurately and robustly. Figure 6 shows the result of hand tracking with depth information. The white point represents the current hand position and the gray points indicate the previous hand positions.

Figure 5 Result of the initial hand detection

4. Conclusion

We proposed a novel hand detection and a tracking method using depth information. We make the motion image, as a basic source of the proposed hand tracking system, which is the accumulated difference image from depth image sequences. In the preprocessing stage, we perform noise reduction, applying a spatial filtering and a morphological processing, and motion clustering,obtaining the moving region from the motion image.

We detect the hand from this motion clusters using waving motion. We also suggest three-axis Kalman filter for tracking. Comparing the proposed method with color-based method, we can see the effectiveness of the proposed method.

Figure 6 Hand tracking using kalman filter

Especially, the depth information based method is very robust to the light variation environment.

As for the future work, in order to improve the accuracy of tracking, more effective noise reduction methods or other tracking methods such as Unscented Kalman filter or particle filter can be considered.

5. References

i. *RG Brown, PYC Hwang, Introduction to Random Signals and Applied Kalman Filtering (Wiley, NY, 1997).*

ii. *C Shan, T Tan, Y Wie, Real-time hand tracking using a mean shift embedded particle filter. Pattern Recogn. 40(7), 1958–1970 (2007). doi:10.1016/j.patcog.2006.12.012.*

iii. *G Bishop, G Welch, An introduction to the Kalman filter, in SIGGRAPH 2001,Course 8, Los Angeles, CA, USA (August 2001).*

iv. *RC Gonzalez, RE Woods, Digital Image Processing, 3rd edn. (Prentice Hall, Upper Saddle River, NJ, 2008).*

v. *I Oikonomidis, N Kyriazis, AA Argyros, Efficient model-based 3D tracking of hand articulations using Kinect, in British Machine Vision Conference, Dundee, UK, pp. 101.1–101.11 (August 2011).*