Reliable and Fast Human Body Tracking under Information Deficiency

Shu-Fai Wong and Kwan-Yee Kenneth Wong
Department of Computer Science and Information Systems
The University of Hong Kong
Pokfulam Rd, Hong Kong
Email: sfwong, kykwong@csis.hku.hk

***Abstract*—** Human body tracking is useful in applications like medical diagnostic, human computer interface, visual surveillance etc. In most cases, only rough position of the target is needed, and blob tracking can be used. The blob region is located within a searching window, which is shifted and resized in each frame based on previous observations. The observations are the locations of the blob in the frames, and are fed into an estimator for predicting the position and the size of the searching window. However, a blob region is regarded as a noisy observation, and the information provided by the blob observation is deficient for most estimators to work well. In this paper, a reliable and efficient estimation algorithm using wavelet is proposed to track human body under information deficiency. The human body is located roughly within a small searching window using color and motion as heuristics. The location and the size of the searching window are estimated using the proposed wavelet estimation scheme. Experimental results show that human body can be tracked accurately and efficiently using the proposed method. The tracker works well in various conditions like clutter background, and background with distractors.

I. Introduction

Human motion tracking is one of the rapidly growing research fields in computer vision. It attracts attention from researchers because of its possible applications in various fields. Motion tracking techniques can be used in various research fields, like video data compression, computer-aided medical diagnosis, virtual reality game development, visual surveillance, and human-computer interface.

Motion tracking involves searching and locating moving object(s) in a video sequence. Commonly used tracking techniques can be found in [1]. The object can be of any shape and can move along any pathway. The number of objects to be tracked is not limited. In most of the tracking algorithms, the object is located by recognition. Recognition can be done based on features (point-to-point matching, correlation analysis on image patch) [2], [3], [4], and object model (object modeling and projection analysis) [5], [6], [7].

Human motion tracking is a kind of motion tracking that applies tracking algorithms on moving human body. Any part of the human body can be tracked. Recent research focuses on tracking of human body parts such as hand, face, and eyes. Hand tracking can be used in sign language recognition. Face tracking can be applied to visual surveillance system. Eyes tracking can be implemented as a kind of human-computer interface.

As in motion tracking, human motion tracking can be done by recognition. In point-to-point matching, eyes and nose on face, and corners like fingertips are extracted in every image of a certain video sequence. The correspondence of these corners in the video frames has to be solved in order to trace the trajectory of the moving human body. In correlation analysis, a small template image of certain human body is generated and is matched against similar patch in every image in video sequence. In object modeling and projection analysis, 3D human body is generated and manipulated such that its 2D projection can match the shape of the detected region in the image frame. A comprehensive survey on human body tracking and analysis can be found in [8], [9].

The tracking of human body is usually slow and inefficient because of the use of recognition. In point-to-point matching, the number of feature points extracted is huge, especially in noisy images. To solve for the correspondence of these points in a different image is time consuming. In correlation analysis, the searching region is large and the image patch comparison is computationally expensive. In object modeling and projection analysis, it takes time to manipulate the 3D model such that it can match the shape of the potential region.

To avoid time-consuming recognition, blob tracking [10], [11] and window searching can be used. Blob tracking involves searching for heuristic features in the image. The heuristic features can be color, motion and shape. Searching for blob features is much faster then searching for predefined object because there is no need to perform detailed recognition. Window searching involves limiting the searching region based on heuristic assumption. The heuristic assumption can be dynamic model of the moving object. Searching within restricted region is much more efficient than searching the whole image.

In addition, blob tracking is appropriate for human body tracking. Human body does contain many heuristic features, e.g. skin color, face configuration, body shape etc. In most applications, there is no need to know the exact position and the pose of the human body. For instance, vision-based human-computer interface and virtual reality game may need the rough position of finger as input. Complicated and time consuming recognition is not neccessary.

Although Blob tracking and window searching is fast and efficient, they are error-prone in most cases. Without detailed recognition, the blob selected may not be the target. This is especially true in the case of noisy image and clutter background. Even though limiting the searching window can facilitate tracking, it is likely that the moving object will move out from the searching window. This is especially true when the dynamic of the moving object is complicated and when the blob or feature extraction scheme is not reliable.

To overcome the accuracy problem of using blob tracking and window searching, reliable feature extraction scheme and estimation algorithms are needed. Feature extraction scheme usually depends on the tracking object. The reliability varies in different environment. For instance, if skin color is extracted as blob feature, the tracking algorithm works well when the environment contains no skin color, but fails when the environment contains clutter background with skin-color distractors. Thus, having a reliable estimation method is very important. By making reasonable estimation based on previous observations, the blob can be tracked even in clutter background and under information deficiency. For example, it is unlikely that an object with downward trajectory will suddenly appear at the top corner of an image.

In this paper, a framework of blob tracking using color model and wavelet-based estimation algorithm will be presented. The human body parts, e.g. hand and face, can be tracked both accurately and efficiently using the proposed system even if the blob observations are of poor quality.

The commonly used estimation algorithms in motion tracking will be reviewed in section II. The proposed architecture will be described in section III. The prediction algorithms and the details of implementation will be illustrated in section IV and V. Finally, research results will be given in section VI.

II. LITERATURE REVIEW

Most efficient motion tracking system involves window searching and motion estimation. By restricting the searching region, tracking can be performed faster. To handle the accuracy problem due to window searching, the location of the window is estimated based on previous observations. Provided that the observations are steady and accurate, the estimation result is usually reliable.

Two commonly used estimation algorithms are Kalman filter [12] and CONDENSATION algorithm [3]. Both algorithms involve the prediction of physical state in next time frame based on previous observations. The internal representation of the physical state in these algorithms is updated through prediction-observation comparison and error-minimization scheme.

In Kalman filter, the prediction is based on the dynamic equation and measurement equation. The dynamic equation describes the transition of physical states. The physical states can be position, velocity, and acceleration of the moving object. Assuming that the object undergoes linear motion, its motion can be described as a transition of internal states and can be formulated by dynamic equation. The transformation from the internal state to observation is done by a measurement equation. The observation should be something that we can measure, e.g. the image location of the object.

Generally, Kalman filter involves a prediction and a updating stages. In the prediction stage, a prediction is made using the dynamic and measurement equations. By applying the dynamic equation to the current physical state, a predicted state is generated. By applying the measurement equation to the predicted state, predicted location is obtained. In the updating stage, the internal state will be updated based on the observation. The prediction is compared with the observation and an error is reported. The Kalman gain is calculated accordingly such that the error can be minimized. The internal states are then adjusted according to the Kalman gain and the process repeats.

The main shortcoming of Kalman filter is its restrictive assumptions. Although Kalman filter is fast in nature, its prediction is made under a lot of strong assumptions. These assumptions include known dynamic model, uni-modal Gaussian noise distribution, non-clutter background, etc. These make Kalman filter fail in various conditions such as non-linear motion tracking, or tracking under noisy condition and clutter background.

To overcome the problem of Kalman filter, CONDENSATION algorithm was proposed. CONDENSATION algorithm stands for conditional density propagation. It aims at finding most probable area containing the feature or object based on sampling. Similar to Kalman filter, CONDENSATION has both the prediction and updating stages. In the prediction stage, estimation is made using the distribution of the samples. For region with high density of samples, the probability of that region contains the target is higher. In the updating stage, observation is used to update the distribution of the sample. Within a searching region, samples are picked up for analysis. For those samples with characteristics that indicate that they belong to the target, they will be considered as candidates. The density of the candidates will be increased according to the updating scheme. By allowing more than one candidate or hypothesis, CONDENSATION algorithm can recover from false tracking of ambiguous feature or object.

Although the accuracy of CONDENSATION algorithm is much higher than that of the Kalman filter, its time complexity is too high. To have high accuracy, the number of sample points should be large. Making inference and updating the density of such a huge number of sample points is time consuming. Furthermore, the candidate regions will converge after long time of updating. By then, CONDENSATION algorithm will have chance of suffering from the false tracking of ambiguous feature.

To track human body efficiently and accuracy under information deficiency, the above two prediction algorithms are perhaps not the candidate choice. In addition to their own shortcomings, they are also too dependent on the observations. These two algorithms base their adjustment purely on the observations. If the observations are not reliable, they will give bizarre tracking result. In the case of blob tracking, the observations are not reliable due to the measurement of rough position. These two algorithms will fail easily under this condition. In order to have fast and reliable tracking using blob tracking techniques, there is a need to have an efficient and reliable estimation scheme.

III. PROSPOSED ARCHITECTURE

Generally, motion tracking involves two steps, namely feature extraction and motion estimation. Within a searching region, features are extracted from the image. To determine whether the features belong to the target, recognition will be performed. After locating the target, the location of the searching window will be shifted or updated based on the results of motion estimation. The features are again extracted from the adjusted searching window again.

To track human body efficiently and accurately, we should have a fast feature extraction scheme and reliable estimation algorithm. As described in section I, feature extraction should not involve precise recognition, which increases time complexity, and estimation algorithm should not suffer from the inaccuracy introduced by feature extraction.

The proposed system has two major components, namely a blob extractor and a wavelet estimator. The blob extractor aims at locating the potential region efficiently using heuristic features. The wavelet estimator aims at predicting the location and the size of the searching window accurately even if the observations provided by the blob extractor is not steady and accurate. The overview of the system is illustrated in figure 1.

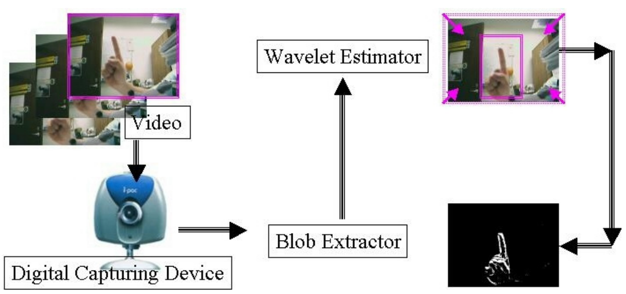

Fig. 1. *The video capturing device will feed input image to the proposed tracking system. Initially, the searching window is set to be the boundary box of the image. Within the searching window, the blob extractor will report the candidate region with change in intensity and skin color. The centroid location and the bounding box of the reported region will be considered as the observation. The wavelet estimator will update the location and the size of the searching window based on the observation. The whole process repeats during tracking.*

A. Blob Extractor

The blob extractor can extract heuristic features from image. The heuristic features can be shape, color, texture or corners. Which features will be used depends on the properties of the target and the tracking environment. In the proposed system, optical field and color model will be used as heuristic features.

Since the human body is moving in the scene, there should be intensity change (optical flow) within the moving region from image to image. For those regions with intensity change, they are considered as candidate region of the moving human body.

Since the human body is of skin color, the region with skin-color can be considered as candidate region. The skin color is represented by a skin-color model [13]. For pixel with color representation closes to the skin color representation in that model, that pixel will be chosen as a potential region.

By combining the optical field and color heuristics, the region with optical flow and skin-color will be selected as the candidate region. The reported region will have high chance of being a region representing the moving human body. Since the reported region is not exactly representing the moving human body, there is a need to have a reliable estimator to avoid inaccurate window adjustment.

B. Wavelet Estimator

The wavelet estimator can make reliable estimation based on noisy observations. As described in section I, the region reported by the blob extractor is rough and unreliable. The extractor may report false region, e.g. a box with skin-color and slight change in intensity. To avoid these kinds of noisy observation, wavelet estimator will ignore the outlying observations and make prediction based on the smoothened trajectory of the target.

Wavelet Estimator performs estimation in two steps, namely trajectory smoothening and trend prediction. In trajectory smoothening step, the input trajectory of the observations is smoothened using wavelet decomposition and reconstruction. The smoothened trajectory will be used to predict the trend using wavelet trend analysis in the prediction stage.

By performing trajectory smoothening, the outlying observations will be removed. The accuracy of the estimation will not be affected even if the observation is not highly reliable. The theoretical background of wavelet smoothening and prediction techniques will be given in section IV.

IV. PREDICTION USING WAVELET

Some of the observations may not be reliable. The blob extractor may report irrelevant regions with heuristic features. Locomotion of human body may introduce the frustration too. This kind of unreliability in observation can be treated as noise. In Kalman filter, this kind of noise is assumed to be uni-modal Gaussian distribution. In real life, however, this kind of assumption is too restrictive. The moving object with locomotion will be ”lost” easily using Kalman filter. For CONDENSATION algorithm, the large sample size will increase the computational time dramatically. Thus, it is not appropriate for real time application.

The proposed wavelet-based estimator can remove the noise of the input trajectory and make prediction based on the trend of the smoothened trajectory. The accuracy of the proposed estimator will not be affected by the outlying and irrelevant observations. The time complexity is not too high such that it can perform quite well even in real time application.

A. Wavelet analysis

The wavelet analysis was initially used in signal processing and financial forecasting [14], [15]. It is widely adopted in these fields because of its accuracy and efficiency in handling curve modeling and smoothening problem [16]. In these applications, observations, e.g. electrical signal, stock price, are usually noisy. In order to make reasonable prediction, we should remove the noise first. Wavelet analysis can decompose the input signal into wavelets for analysis. By removing those wavelets representing noise, the output curve is smoothened. Further analysis and prediction can be made on the smoothened signal.

Wavelet analysis consists of two steps, namely decomposition and reconstruction. In wavelet decomposition, input signal is broken down into small components, called wavelets. The wavelets have internal parameters, such as scale and transition. The mother wavelet is represented by a certain function, e.g. Harr function. The wavelets are differentiated from each other in terms of their scale and transition, but not the fundamental shape. By breaking down signal into wavelets, a spectrum is formed. The spectrum is in two dimensions, the scale and transition. Every signal has its own spectrum. Analysis can be performed on this spectrum, e.g. finding out the major components of the signal. Given a wavelet spectrum, signal can be reconstructed from it. By performing wave superposition operation on the wavelets according to the wavelet coefficient (i.e. the intensity in wavelet spectrum), original signal can be reconstructed. The formula for wavelet decomposition and reconstruction is given by:

$$\phi_{j,k}(t) = 2^{-\frac{j}{2}}\phi(2^{-j}t-k), \psi_{j,k}(t) = 2^{-\frac{j}{2}}\psi(2^{-j}t-k) \quad (1)$$

$$f(t) = \sum_{k\in Z} a_k^J \phi_{J,k}(t) + \sum_{j\leq J}\sum_{k\in Z} d_k^j \psi_{j,k}(t) \quad (2)$$

where f(t) is the input signal at time t, ϕ is the scaling function, ψ is the mother wavelet, 2^{-j} and k represent the scaling and transition factor, a_k^J and d_k^j are the scaling and wavelet coefficients respectively.

B. Denoising using wavelet analysis

Denoising can be done by wavelet decomposition and selective reconstruction. By decomposing the input trajectory (the input signal) into the constituent wavelets, the major wavelets can be identified from the wavelet spectrum. The wavelet spectrum is shown in figure 2. Re-combination of these major wavelets forms the smoothened trajectory. The minor wavelets represent the noise and are removed. The decomposition and reconstruction are illustrated in figure 3.

C. Prediction using wavelet analysis

Prediction can be done by wavelets superposition. Combining the constituent wavelets near to the current time frame and calculating the value of superposition will give the estimated location at next time frame. The idea is illustrated in figure 3.

V. ALGORITHMS AND IMPLEMENTATIONS

As described in section III, the proposed system consists of the blob extractor and the wavelet estimator. The blob extractor can extract heuristic feature quickly but with noise and not reliable. The wavelet estimator can denoise the output from

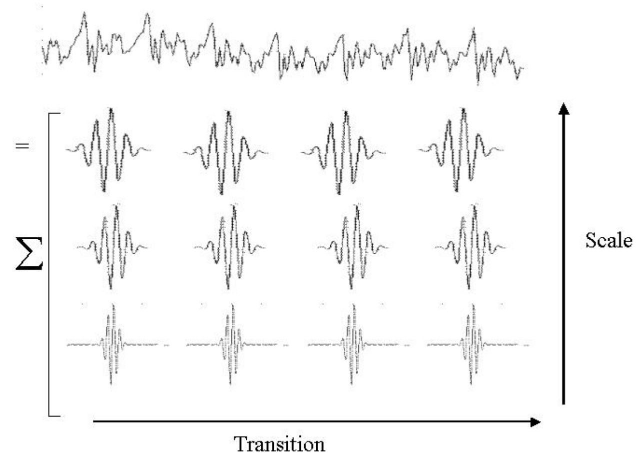

Fig. 2. *The input signal can be decomposed into its constituent wavelets. All these wavelets can be represented by the same mother wavelet function, and thus they have similar shape. Each wavelet has two parameters, the scale and transition, which make them different from one another. Wavelet spectrum of the input signal represents an array of wavelet coefficients of the constituent wavelets along scale and transition dimension. Every signal has its own spectrum. Using the spectrum, original signal can be reconstructed.*

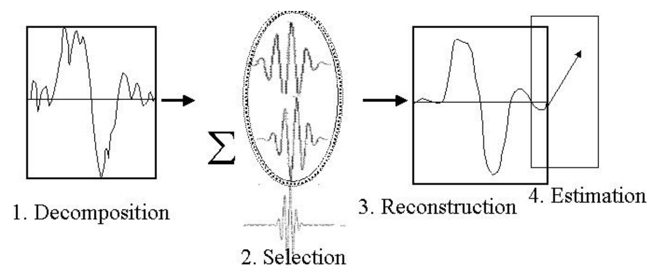

Fig. 3. *The input signal can be considered as the superposition of wavelets with different scales and transitions. Major wavelets are selected among the pool of constituent wavelets. Smoothened signal can be reconstructed from those major wavelets. Combining the major wavelets near current time frame will give prediction result of the next time frame.*

the blob extractor and make prediction on the location of the searching window. The implementation details are given in following sections.

A. Blob extractor

In the proposed system, the blob extractor aims at locating the moving human body. We can make use of some characteristics of moving human body to facilitate the target searching. Optical flow and skin color are selected as the heuristic features in this system.

The Optical flow or change in intensity is detected by reference white adjustment and background subtraction. Due to the variation of lighting, especially under fluorescent light, reference white adjustment have to be performed to reduce this variation. The change in intensity is then detected by background subtraction and thresholding. The intensity change detected between images may be due to the motion of the object. This region is selected as a candidate for further investigation.

The skin color region is extracted based on the skin color model. This model has been widely used in face detection research, e.g. [13]. The color intensity of every pixel within the searching region is transformed into the YC_bC_r coordinates. Its distance from the coordinates of possible skin color is used to determine whether that pixel is of skin color.

By using only one heuristic feature, the candidate region reported is not accurate. Thus, the final reported region is the intersection set of both regions with change in intensity and skin color. The result is shown in figure 4.

B. *Wavelet estimator*

In the proposed system, the wavelet estimator aims at adjusting the location and the size of the searching window based on previous observations. The observation refers to the centroid of the reported region of the blob extractor. The location of centroid is broken down into the x and y direction. The measurements of the x and y components will then be fed into 2 wavelet estimators separately. The measurements of the x and y components in a certain range of time frames form the input signal for the estimators. However, the observations may be noisy due to inaccuracy of the blob extractor, clutter background and locomotion of hand. Denoising will be done to smoothen the trajectory of the input signal. Prediction will be performed on the smoothened signal afterwards. The prediction of the estimators will form the coordinate of the predicted location of the searching window. The prediction error will be used to adjust the size of the searching window. The higher the prediction error, the larger the size of the searching window will be.

Denoising is done by wavelet decomposition and selective reconstruction. The input signal is broken down into its constituent wavelets. Those wavelets with large scale (low frequency) parameter and with large wavelet coefficient will be selected. To remove ripples and noisy signal from the input trajectory, wavelets with low frequency is selected. To remove irrelevant and weak interfering signal, only wavelets with high coefficient is selected. Reconstructing the signal using wavelets with low frequency and high coefficient values gives the smoothened trajectory.

Prediction is done by wavelet superposition operation. The trend of the signal can be inferred from the wavelets with low frequency and high coefficient value. The estimation can be formulated as:

$$Est(t+1) = \sum_{i=1}^{N} w_i \psi_i(s_i, k_i, t+1) + w_0 Mean \qquad (3)$$

where Est(t) is the estimation made at time t+1, N is the number of major wavelets, ψ_i are the major wavelets, w_i are the corresponding wavelet coefficient and s_i, k_i are the scaling and transition parameters of that major wavelet.

It represents the value of the superposition of the major wavelets at the prediction time frame. The result is based on the trend of the smoothened input trajectory.

VI. EXPERIMENTAL RESULT

The proposed system was implemented using Visual C++ under Microsoft Windows. Three experiments was performed to test the system. The experiments were done on a P4 2.26 GHz computer with 512M ram running Microsoft Windows.

A. *Experiment 1: 1-D signal analysis*

In this experiment, the smoothening and prediction performance of the wavelet estimator was tested. The input signal was the y-component of the 2D moving path of the hand. The result is shown in figure 5 and 6. The result shows that the wavelet estimator can handle noisy signal quite well. The smoothened signal is close to the actual movement and the prediction error is low.

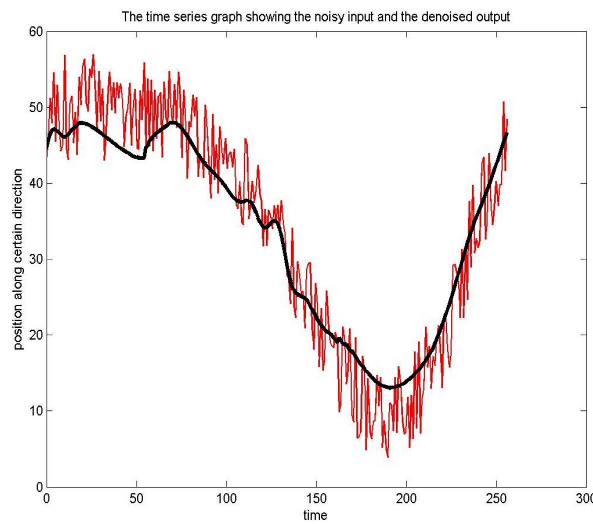

Fig. 5. *The red fuzzy line represents the input signal (the y-component of the 2D moving path of the hand). The black solid line represents the smoothened signal using wavelet analysis. Even the input signal contains noise due to unreliability of blob extraction, the trajectory can be still smoothened using wavelet denoising.*

B. *Experiment 2: moving faces*

In this experiment, the face under clutter background was tracked. The result is shown in figure 7. It shows that the face can be tracked even when the background consists of skin color.

C. *Experiment 3: moving hand*

The hand under clutter background was tracked in this experiment. The result is shown in figure 8. Similar to experiment 2, the hand can be tracked in the clutter background consists of skin color. In addition, when there is locomotion of hand, the trajectory can still be smoothened.

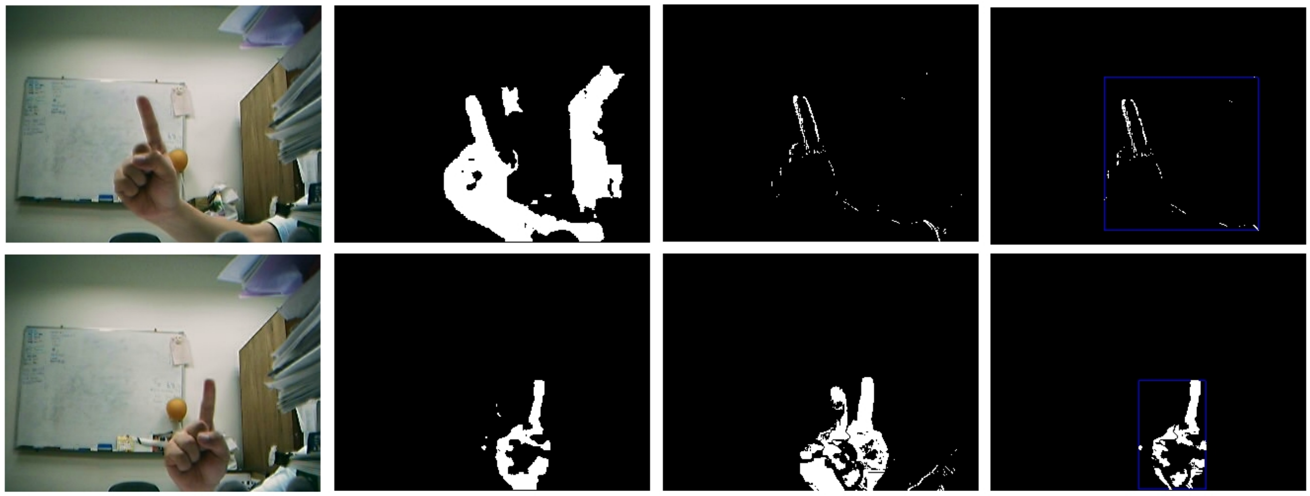

Fig. 4. *These two rows show the output images of applying different blob selection schemes. The left most images are the input images. The images in the second column are the output of using color detector only. The images in the third column are the output of applying optical flow extraction only. The right most images show the output by using both color and moving region detectors. It shows that using both color and moving region as heuristic features is much more reliable than using any one of them alone.*

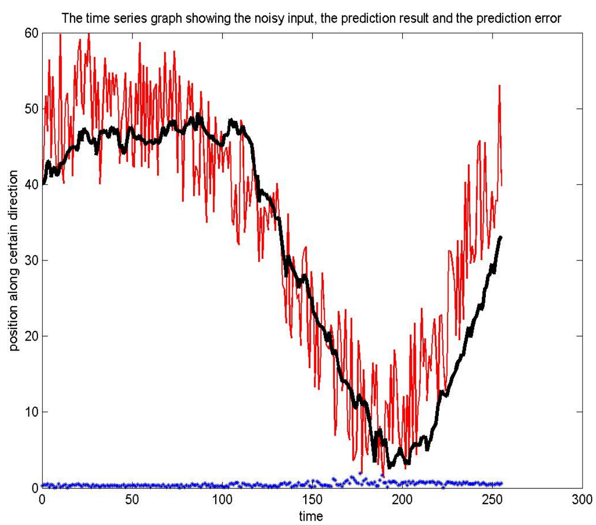

Fig. 6. *The red fuzzy line represents the observation (the y-component of the 2D location of the hand) at each time frame. The black solid line represents the prediction at each time frame using wavelet prediction approach. The blue marker line at the bottom shows the relative prediction error. It shows that prediction using wavelet is reliable even if the signal is very noisy. The predictions are usually close to the observations.*

VII. Conclusions

This paper endeavors to look for a way to track human body efficiently while maintaining reasonable high accuracy. Blob tracking, instead of time-consuming recognition approach, is adopted. By limiting the size of the searching window, the blob region can be located much faster. To keep track of the target, the searching window should be shifted and resized according to the observations. However, the noise introduced by blob tracking will decrease the accuracy of estimators seriously. The commonly used Kalman filter and CONDENSATION algorithm perhaps cannot solve the problem efficiently and accurately. This paper proposes the wavelet estimator that can perform a relatively reliable estimation efficiently even under information deficient and noisy environment. By using the proposed system, human body can be tracked by blob tracking techniques accurately and efficiently.

In the proposed system, it uses skin color model, and optical field to locate the rough position of the human body (i.e. the blob region) within a searching window. The searching window is automatically shifted and resized according to the prediction made by the wavelet estimator. Tracking process is fast because of the rough scale blob searching and efficient estimation scheme for the searching window. The accuracy of the estimation remains high even when the object location is rough and the information is not reliable.

We have tested the performance of the proposed tracking system using 3 sets of experiments. In the first experiment, an 1D signal was analyzed using the wavelet estimator. It shows that the noisy signal can be smoothened and the predictive power of the estimator is quite high. In the remaining two experiments, the moving human parts were tracked in real time. The results show that the human body (e.g. hand and face) can be tracked even in clutter background consisting of skin color, and the locomotion of the object will not affect the tracking result seriously.

Although the proposed tracker can give relatively reliable estimation efficiently, it can only track a blob region instead of an exact location of the target. If we wish to have a better and more exact location of the whole object, the computational time will increase dramatically. Modification and improvement have to be done on the object location algorithm in order to have more accurate and robust tracking result.

Fig. 7. *The green (lighter) block is the blob region detected. The blue (darker) block is the maximum bound of the searching window. The red fuzzy dots are the observations (the centroid of blob region). The green continuous line is the predicted trajectory which describes the smoothened trend of the blob region. It shows that the face can be tracked accurately even under clutter background with distracting color (skin color of the balloon, bookshelf etc.).*

VIII. REFERENCES

[1] A. Blake and A. Yuille, *Active Vision*. Cambridge, MA: MIT Press, 1992.

[2] M. Isard and A. Blake, "Contour tracking by stochastic propagation of conditional density," in *European Conference on Computer Vision*, 1996, pp. 343–356.

[3] ——, "Condensation – conditional density propagation for visual tracking," *International Journal of Computer Vision*, vol. 29, no. 1, pp. 5–28, 1998.

[4] ——, "ICONDENSATION: Unifying low-level and high-level tracking in a stochastic framework," *Lecture Notes in Computer Science*, vol. 1406, pp. 893–908, 1998.

[5] M. J. Black and A. D. Jepson, "Eigentracking: Robust matching and tracking of articulated objects using a view-based representation," in *European Conference on Computer Vision*, 1996, pp. 329–342.

[6] D. DeCarlo and D. N. Metaxas, "Optical flow constraints on deformable models with applications to face tracking," *International Journal of Computer Vision*, vol. 38, no. 2, pp. 99–127, 2000.

[7] V. Kruger, A. Happe, and G. Sommer, "Affine real-time face tracking using a wavelet network," in *Proc. Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems*, 1999, pp. 141–148.

[8] D. M. Gavrila, "The visual analysis of human movement: A survey," *Computer Vision and Image Understanding: CVIU*, vol. 73, no. 1, pp. 82–98, 1999.

[9] T. B. Moeslund and E. Granum, "A survey of computer vision-based human motion capture," *Computer Vision and Image Understanding: CVIU*, vol. 81, no. 3, pp. 231–268, 2001.

[10] A. Azarbayejani and A. Pentland, "Real-time self-calibrating stereo person tracking using 3-d shape estimation from blob features," *Tech. Report, M.I.T Media Laboratory Perceptual Computing Section*, no. 363, 1996.

[11] C. R. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: Real-time tracking of the human body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780–785, 1997.

[12] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[13] R. Hsu, M. Abdel-Mottaleb, and A. Jain, "Face detection in color images," in *Proc. IEEE ICIP*, 2001, pp. 1046–1049.

[14] C. Chang, W. Fu, and M. Yi, "Short term load forecasting using wavelet networks," *Engineering Intelligent Systems for Electrical Engineering and Communications*, vol. 6, pp. 217–2223, 1998.

[15] D. B. Percival and A. t. Walden, *Wavelet Methods for Time Series Analysis*. Cambridge: Cambridge University Press, 2000.

Fig. 8. *The green (lighter) block is the blob region detected. The blue (darker) block is the maximum bound of the searching window. The red fuzzy dots are the observations. The green continuous line is the predicted trajectory which describes the movement of the blob region. It shows that the hand can be tracked accurately under clutter background as in face tracking. In addition, even when there is locomotion of the hand, the predicted trajectory of the blob was not affected much.*

[16] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia: SIAM, 1992.