Virtual Input Devices for Underwater Manipulator

Nadira Nordin, Umair Soori and Mohd Rizal Arshad

USM Robotics Research Group, School of Electrical and Elctronic Engineering,
Universiti Sains Malaysia, Engineering Campus,
14300 Nibong Tebal, Pulau Pinang, Malaysia
Email: n_adira@yahoo.com, umsoori@yahoo.com, rizal@eng.usm.my

ABSTRACT

Computer vision gives the potential to present a vision-based hand tracking system for virtual input. Tracking hand movement has become popular, efficient and more natural in human computer interaction. However available systems are invasive and require the user to wear gloves or markers. In this paper we present, a non-invasive hand tracking system as virtual input devices for underwater manipulator. In this system, coordinate of the palm represent as the input for the manipulator. So the main objective to the system is to track the palm and get the actual coordinate in 3D workspace. As the system is a non-invasive system, there is no markers attach to the palm. So in this paper, we will discuss on the palm segmentation using skin color detection technique and how to utilize probabilistic estimation, Kalman filter algorithm based on Bayesian framework which estimates motion and direction of moving palm.

1 INTRODUCTION

Recently, Human-Computer Interaction (HCI) becomes popular in everyday applications. It has potential in home and industrial applications such as visual surveillance, gait analysis and gesture analysis. Traditionally, HCI input devices is based on mechanical devices such as mouse, keyboard, joystick and gamepads but nowadays, computer vision (CV) has the potential to provide natural, non-contact solutions [1]. Means using a camera, CV will capture and translate user's movement such as motion of head, eye gaze, face, hand, arms or even the whole body into semantics indicatory symbols to manipulate machine operation.

In this paper, we will discuss more on hand-based input for HCI. At present, marker-based hand motion capture is being used such as data gloves magnetic sensing devices [2]. Data glove is an interactive device, similar to a glove worn on the hand, which facilitates magnetic sensing and fine-motion control to calculate the finger joint angles. A motion tracker also attached to capture the global position or rotation data of the glove. This device can give very accurate result, however the cost for hardware is expensive, invasive and require complex setup procedures.

A marker less system using CV represents a promising alternative to marker-based hand motion capture because it is non-invasive, easy to setup and natural. A marker less system of human motion capture could be run using conventional cameras and without the use of special apparel or other equipment. In this paper, we propose a real-time non-invasive hand tracking and plot the actual coordinate in 3D graph for underwater manipulator.

This paper is organized as follows. Section 2 contains palm segmentation using skin color detection technique. This followed by utilize probabilistic estimation, Kalman filter algorithm based on Bayesian framework which estimates motion and direction of moving palm in Section 3. The experimental results and 3D plotting are presented in Section 4. The last section concludes the paper. The conceptual block diagram of the algorithm is depicted in Figure 1.

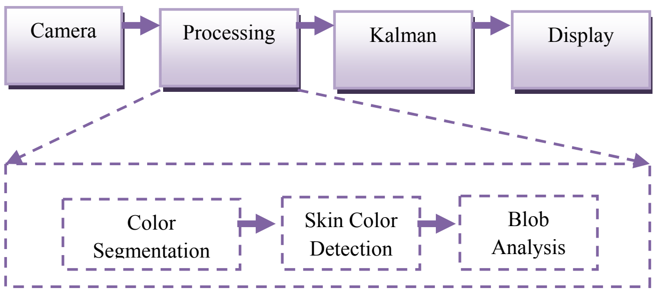

Figure 1: Conceptual block diagram of the proposed algorithm.

2 IMPLEMENTATION

In our application, coordinate of the palm represent as the input for the end effecter of the underwater manipulator. The initial step is palm segmentation where the image region that contains the hand has to locate. In order to make this process, we can use shapes, but they vary greatly during natural motion of hand [3]. Therefore, we choose skin colour as the palm feature because it allows fast processing and is highly robust to geometric variation of the palm pattern [4].

2.1 Color Segmentation

RGB is combination of three colours ray (red, green and blue) and most popular colour spaces for processing and storing of digital image data. However, high correlation between channel, significant perceptual non-uniformity, mixing of chrominance and luminance data make RGB not very encouraging choice for color analysis and color based recognition algorithms[4].

YC_rC_b colour space is choose because it is an represented by *luma* (which is luminance, computed from nonlinear RGB, constructed as a weighted sum of the RGB values, and two color difference values C_r and C_b that are formed by subtracting *luma* from RGB red and blue components.

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ C_r &= R - Y \\ C_b &= B - Y \end{aligned} \qquad (1)$$

Because of the transformation simplicity and explicit separation of luminance and chrominance components, this colour space is suitable for skin color modeling [5], [6], [7].

Human skin colours can be represented by an elliptical Gaussian joint probability density function (pdf), defined as:

$$p[\underline{x}(i,j)/W_s] = (2\pi)^{-1} |\Sigma_s|^{-1/2} \exp[-\lambda^2_s(i,j)/2] \qquad (2)$$

Where $\underline{x}(i,j) = [\underline{x}(i,j)\underline{y}(i,j)]^T$ represents the random measured values of the chrominance (x,y) of a pixel with coordinates (i,j) in an image, W_s is the class describing skin, Σ_s is covariance matrix for the skin chrominance, and where $\lambda_s(i,j)$ is the Mahalanobis distance from vector $\underline{x}(i,j)$ to mean vector $\mu s = [\mu x_s \mu y_s]^T$ obtained for skin chrominance, defined as

$$[\lambda_s(i,j)]^2 = [x(i,j) - \mu_s]^T \Sigma_s^{-1} [x(i,j) - \mu_s] \qquad (3)$$

Mahalanobis distance from the color vector $\mathbf{x}(i,j)$, to mean vector μ_s, given the covariance matrix Σ_s can serve for the same purpose[8]. Figure 2 shows skin color segmentation.

Original Image Segmented Image

Figure 2: Segmentation of the skin color from original image.

2.2 Blob Analysis

Blob is defined as a region of connected pixels, while blob analysis is the identification and study of these regions in an image. The algorithms discern pixels by their value and place them in the image [9]. The blob features usually calculated area and perimeter, Ferret diameter, blob shape, and location. The flexibility of blob analysis is, it classifies image pixels as object by some means, joins the classified pixels to make discrete objects using neighborhood connectivity rules, and computes various properties of the connected objects to determine position, size, and orientation. In this paper, the bounding blob is the skin segmentation, so analysis tools typically consider touching foreground pixels to be part of the same blob. We named this moving block as a motion block, which is tracking the palm performing motion as a shape of rectangle. Using the blob's bounding box, we can alleviates the need to search for corresponding point features in consecutive frames [9]. See Figure 3. The point or centroid values will represent the spatial coordinate locations of the palm.

Figure 3: Blob analysis at the palm area.

2.3 Probabilistic Estimation

Detecting the palm using the skin colour segmentation and using the blob analysis to keep track the palm movement is not enough. Sometimes the detection will lost due to the speed and noise. So we used Bayesian estimation approach to keep tracking the hand without losing any data.

The proposed trajectory estimation scheme is quite simple and we have assumed that the palm motion is a linear motion. Bayesian estimation, a probabilistic approach is used for estimating the uncertain probability density function over time. R.E. Kalman presented a novel approach in 1960. The Kalman filter is a recursive estimator. It is a tool that can estimate the variables of a wide range of processes. In mathematical terms Kalman filter estimates the states of a linear system [10].

The set of recursive equations as describe below, known as the Kalman filter, consists of two repeating stages, prediction and updating in the presence of process noise. Brief review of the equations of Kalman Filter algorithm as follow:

- ***Predicted equation***

$$x(k) = Ax(k-1) + Bu(k-1) + W(k-1) \quad (4)$$

- ***Measurement equation***

$$z(k) = Hx(k) + V(k) \quad (5)$$

Where ***x (k)*** and ***z (k)*** are state and measurement vector at time ***k***. ***A***, ***B*** and ***H*** are state transition matrixes. The ***W (k), V (k)*** variables and represent the process and measurement noise respectively, probability distribution of ***W (k)~N (0, Q)*** and ***V (k)~N(0,R)***

Time Update ("Predict")

- ***State prediction***

$$\hat{x}^{-}(k) = A\hat{x}(k-1) + Bu(k) \quad (6)$$

- ***Prediction-error covariance***

$$P^{-}(k) = AP(k-1)A^{T} + Q \quad (7)$$

Initial estimates for $\hat{x}_{(k-1)} = 0$ and $P_{(k-1)} = 0$

Measurement Update ("Correct")

- ***State updating***

$$\hat{x}(k) = \hat{x}^{-}(k) + K(k)(z(k) - H\hat{x}^{-}(k)) \quad (8)$$

- ***Kalman gain***

$$K(k) = P^{-}(k)H^{T}(HP^{-}(k)H^{T} + R)^{-1} \quad (9)$$

- ***Updating-error covariance***

$$P(k) = (I - K(k)H)P^{-}(k) \quad (10)$$

The superscripts ***'-"***denote ***''before''*** the measurement and ***I*** denoted as Identity matrix. In our work Kalman filter estimates ***t+1*** frame information of moving object that is motion block, as the method depend upon real-time tracking so in this application we are restricted to the real estimation ***t+1*** at time ***t***. If we used Kalman filter to estimate further more frames other then ***t+1***then we would lost the real tracking of the object at time ***t***. The process noise covariance ***Q*** in the time delay between two consecutive frames, and measurement noise covariance ***R*** is the error, i.e. difference between measure and estimated observations in pervious iteration.

The transition matrixes have been taken from 2D example of position and velocity [11] where

State Transition

$$A = \begin{bmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Measurement matrix

$$\boldsymbol{H} = \begin{bmatrix} H_x & 0 & 0 & 0 \\ 0 & H_y & 0 & 0 \end{bmatrix}$$

Where $\boldsymbol{dt}, \boldsymbol{H_x}$ ***and*** $\boldsymbol{H_y} = 1$

3 RESULTS AND DISCUSSION

Tracking the palm without any markers is quite challenging due to the environment and lighting. We had conducted a experiment to demonstrate the palm movement tracking. The experiment was conducted in a normal lab with ceiling lighting. This test implements the colour segmentation and probabilistic estimation in MatLab environment and all test have been conducted on a 1.80 GHz core2duo Pentium IV computer, executing Windows XP.

The algorithm will search for the palm by using the skin color segmentation. Once the palm has been detected, the system will display in binary result to differentiate between the palm and the background. To get better result, we perform dilation operation followed by an erosion operation using predefined neighborhood or structuring element. A segmented window will show the result.

Then the system performs the blob analysis to the segmented image and we can see the bounding blob only cover the palm area. From that we can get the centroid of the palm that represents the actual coordinate for the end effecter of the underwater manipulator. Kalman filter is used to estimate motion and direction of moving palm. Hence the tracking data will not lost during the tracking. As the result, we can see the trajectory of the palm movement in 3D plotting graph. See Figure 4.

4 CONCLUSION

In this paper, a non-invasive palm tracking system is presented. The developed vision system successfully track the skin-color of the palm which can be easily setup compare to other invasive system like data glove. Using the blob analysis, we can easily get the centroid of the palm and we used the centroid as the actual coordinate for the underwater manipulator. The plotting graph shows the trajectory of the moving palm in 3D space.

5 ACKNOWLEDGEMENTS

This part of research is funded by Malaysian government, Ministry Of Science, Technology and Innovation (MOSTI) Grant.

6 REFERENCES

[1] Ali Erol, George Bebis, Mircea Nicolescu, Richard D. Boyle, Xander Twombly "Vision-based hand pose estimation: A review", *Computer Vision and Image Understanding* 108 (2007) 52-73.

[2] David J. Sturman, David Zeltzer, "A Survey of Glove-Based Input", *IEEE Computer Graphics & Applications* (1994).

[3] T. Heap, D. Hogg, "Wormholes in shape space: tracking through discontinuous changes in shape", *Proc. Sixth International Conference on Computer Vision, ICCV98*, pp. 344-349, 1998.

[4] Vladimir Vezhnevets, Vassili Sazonov and Alla Andreeva, "A Survey on Pixel-Based Skin Color Detection Techniques, *in Proc. Graphion-2003,85-92.*

[5] Phung, S. L., Bouzerdoum, A, and Chai, D., "A novel skin color model in ycbcr color space and its application to human face detection". In *IEEE International Conference on Image Processing (ICIP'2002)*, vol. 1, 289–292.

[6] Ahlberg, J., "A system for face localization and facial feature extraction", *Tech. Rep. LiTH-ISY-R-2172, Linkoping University*, 1999.

[7] Chai, D., and Bouzerdoum, A., "Bayesian Approach to Skin Color Classification in ycbcr Color Space", In *Proceedings IEEE Region Ten Conference (TENCON'2000)*, vol. 2, 421–424.

[8] Terrillon, J.-C, Shirazi, M. N., Fukamachi, H., and Akamatsu, S., "Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color

images”, In *Proc. of the International Conference on Face and Gesture Recognition* (2000), 54–61.

[9] Romer Rosales and Stan Sclaroff, “3D Trajectory Recovery for Tracking Multiple Objects and Trajectory Guided Recognition of Actions”, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, June 1999, pp. 123 Vol. 2.

[10] Muhammad Asif and Mohd. Rizal Arshad 2007 “AGV guidance system: an application of active contour and kalman filter for road tracking”, *Proceedings of the 4th International Symposium on Mechatronics and its Applications (ISMA07)*, Sharjah, U.A.E. March 26-29, 2007.

[11] Greg Welch and Gary Bishop, *An Introduction to the Kalman Filter*, ACM, Inc. 2001.

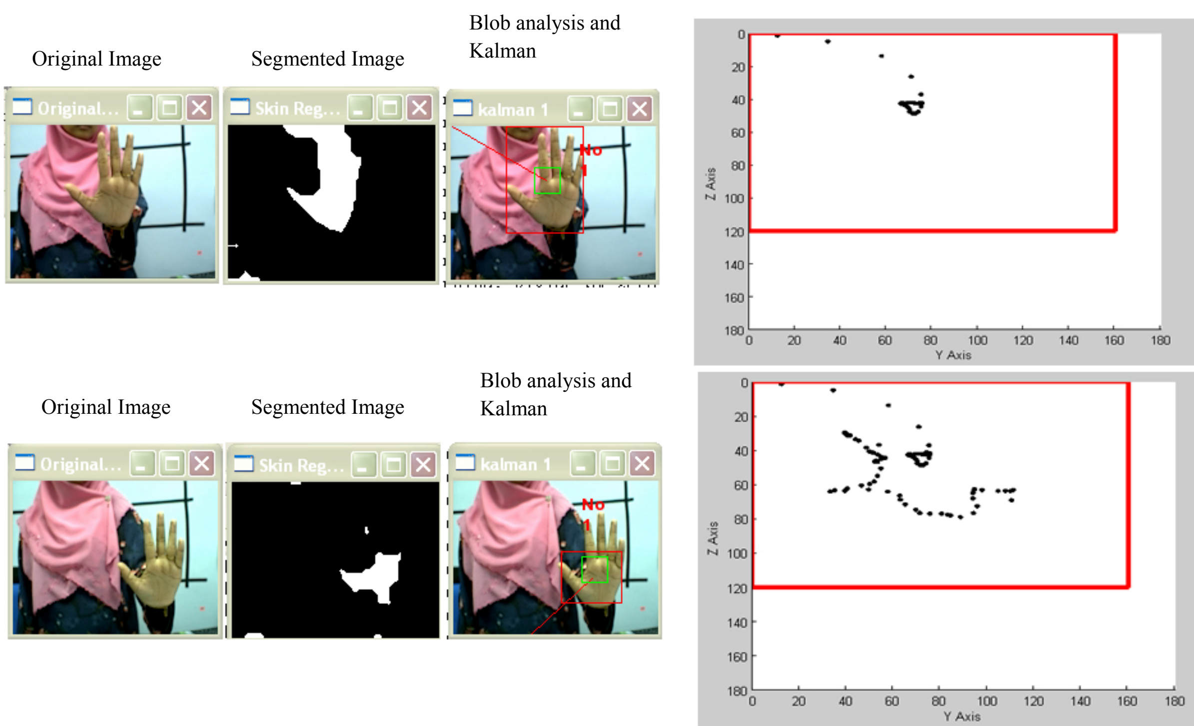

Figure 4: The consequence result from the system algorithm. The estimation shows in blob analysis and kalman image where Green blocks is Kalman prediction, Red blocks real detection and line shows estimated direction. At the right side side the plotting graph during palm tracking.