A Study for Vision Based Data Glove Considering Hidden Fingertip with Self-Occlusion

Sanshiro Yamamoto*
Dept. of Computer Science and Eng.
Nagoya Institute of Technology
Nagoya, Japan

Kenji Funahashi
Dept. of Computer Science and Eng.
Nagoya Institute of Technology
Nagoya, Japan
Email: kenji@nitech.ac.jp

Yuji Iwahori
Dept. of Computer Science
Chubu University
Kasugai, Japan

***Abstract*—Data glove is widely used interface device which measures hand posture (finger joint angles) and inputs it into a computer in virtual reality field. But it dose not spread throughout home because of expensive interface. On the other hand, researches to recognize a hand posture from photo/video images are performed, which are a kind of hand posture measurement system and called Vision Based Data Glove (VBDG) in recent years. In this paper, we propose a new VBDG system. It estimates hand motion using detected fingertip positions with monocular camera and inverse kinematics. However it cannot estimate hand motion when a fingertip is undetectable with self-occlusion. So our system estimates hidden finger motion, then estimates hand motion. Our experimental results show that the proposed method can estimate it with sufficient accuracy in real-time. Since camera base system is inexpensive, it fits personal use.**

***Keywords*-Vision Based Data Glove, Self-Occlusion, Hidden Fingertip, Inverse Kinematics, Personal Use**

I. Introduction

Data glove is widely used interface device which measures hand posture (finger joint angles) and inputs it into a computer in virtual reality field. However, high-performance data glove is expensive and low-priced data glove does not have enough sensors to capture hand data correctly. It dose not spread throughout home. On the other hand, researches to recognize a hand posture from photo/video images are performed, which are a kind of hand posture measurement system and called Vision Based Data Glove (VBDG) in recent years. VBDG system can be classified into two groups. One is use database of hand postures built in advance, it estimates hand posture to match an input image with database. Another detects fingertips from an input image first, then it recognizes hand gesture and/or estimates hand posture.

The hand posture estimation method to match an input image with database [1,2] can be registered various hand posture, so it can cope with self-occlusion. However a flexibility of human hand is so high and so many postures can be targets to match with database. Therefore process speed and estimating accuracy are trade-offs. The method to search efficiently is also proposed [3]. Although using these methods, much data of hand posture is required, it has wide individual difference probably. So it is difficult to build database for each user at home.

On the other hand, the method detecting fingertips can be further classified into two methods. These are a hand gesture recognition method from detected fingertips [4,5] and a hand posture estimation method which solves Inverse Kinematics (IK) with fingertip positions [6,7,8]. First one recognizes a hand gesture with combination of detected fingertip motion. But this does not estimate each finger joint angle. It is not so useful for general application, i.e. to realize many types of object operation. On second one, hand model information (hand size, finger length, etc.) is required to solve IK although, to create hand model is easier than to build database for each user at home. However it cannot estimate hand posture when a fingertip is undetectable with self-occlusion. Chen et al. proposed to choose suitable images from several cameras and to estimate it [7]. It is difficult to set several cameras at home.

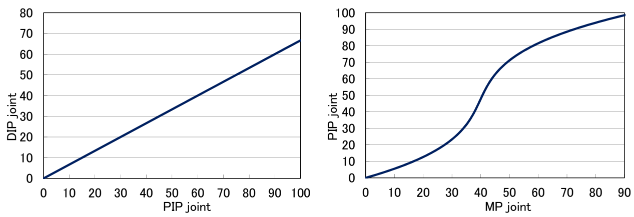
In this paper, we propose a hand posture estimation method after fingertip detection using computer vision technology, even when a camera cannot detect some fingertips with self-occlusion. Our experiment system is based on our pilot VBDG system without self-occlusion [9].

In proposed method, we assume that the finger motion is continued for a while when a finger is hidden. Since hand moves smoothly, we also assume that finger motion (time series variation of each finger joint angle) can be expressed with polynomials. The finger posture in current frame is estimated using the polynomials first. However this estimation is not necessarily correct. So we also consider the finger motion of the hidden fingertip. There we call the finger of hidden fingertip as hidden finger. Both of fingertip positions of estimated finger posture and previous frame posture are judged whether these are hidden with self-occlusion. Then hidden finger motion type is estimated and hidden finger posture is estimated where the fingertip is hidden with self-occlusion. Then our system estimates hand

*present affiliation is HAGIWARA ELECTRIC Co.,Ltd.

Table I: Movable Range of Finger Angle

	IP	MP	CM	CM-AA
THUMB	$0° - 80°$	$0° - 60°$	$0° - 90°$	$0 - 60°$
	DIP	PIP	MP	MP-AA
INDEX	$0° - 80°$	$0° - 100°$	$0° - 90°$	$-25° - 25°$
MIDDLE	$0° - 80°$	$0° - 100°$	$0° - 90°$	$-25° - 25°$
RING	$0° - 80°$	$0° - 100°$	$0° - 90°$	$-25° - 25°$
LITTLE	$0° - 80°$	$0° - 100°$	$0° - 90°$	$-25° - 25°$



(a) (b)

Figure 1: Correlation of finger joints. The figure (a) shows relation between PIP and DIP joint. The figure (b) shows relation between MP and PIP joint.

posture solving IK from fingertip positions when a fingertip is undetectable with self-occlusion.

II. Pilot VBDG System

First of all, we would like to assume that fingertips are already detected from an input 2D image with computer vision technology. So our pilot VBDG system detect them using simple and easy method as followed. Fingertips are detected using color markers attached on the fingertips, and the hand position as 6 degrees of freedom is measured with an ARToolKit marker [10] attached on the palm. Then Fingertip 3D positions are calculated considering hand structural constrains. The hand posture (finger joint angles) is estimated with IK.

A human hand has many constraints. For example, first, the movement of each finger is limited with movable range of finger angle as shown in Table I. DIP is distal interphalangeal joint, PIP is proximal interphalangeal joint, IP is interphalangeal joint, MP is metacarpophalangeal joint, CM is carpometacarpal joint and AA is abduction and adduction. Second, finger joint angles have correlation each other as shown in Figure 1. There are many constraints described above. If a finger moves following as these constraints, a fingertip moves on a defined curve. This pilot system estimates fingertip positions using this curve.

It solves IK with Cyclic Coordinate Descent method [11], then gets finger joint angles. This CCD method calculates them solving optimization problem which minimizes the distance between an end effector and the goal. Our experiment system is built based on this pilot system.

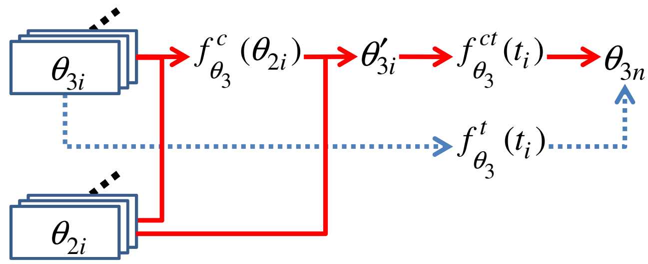


Figure 2: Derivation of a cubic polynomial showing time series change of MP joint.

III. Hidden Finger Posture Estimation

A. Hidden Finger Posture Continuing Finger Motion

Hand posture changes continuously. We assume that a hidden finger motion is continued for a while when a finger is hidden. A hidden finger posture can be estimated from the hidden finger motion (transition of each finger joint angle in past frames). Moreover, the hand moves smoothly. Transition of each finger joint angle in past frames can be approximated to a cubic polynomial. Therefore, it is possible that a finger motion in past frames has correlation as shown in Figure 1. The correlation in past frames is derived. And the cubic polynomial showing time series change which includes correlation is derived. When MP-AA joint data without the correlation which becomes settled uniquely, or the past frames data is not mutually related from other finger joint angles, a cubic polynomial is derived only from transition of the finger joint angle in past frames. A cubic polynomial is derived using least-squares method.

I will show an example to illustrate derivation of a cubic polynomial showing time series change of MP joint. Figure 2 shows the cubic polynomial derivation showing time series change of MP joint, where θ_{3i} is MP joint angle, θ_{2i} is PIP joint angle. i is frame number ($p \leq i \leq n$), p is oldest frame to refer to past frames, now n is current frame and t_i is lapsed time to i frame when $t_p = 0.0$. Polynomial $f^c_{\theta_3}(\theta_{2i})$ showing the correlation of θ_{3i} and θ_{2i} is derived from past frame data of each joint. Angle θ'_{3i} in past frames which includes correlation is generated from the polynomial $f^c_{\theta_3}(\theta_{2i})$ and θ_{2i}. Polynomial $f^{ct}_{\theta_3}(t_i)$ is derived from θ'_{3i} and t_i. Then θ_{3n} is obtained with giving t_n to the polynomial $f^{ct}_{\theta_3}(t_i)$. When the finger motion is continued for a while after the fingertip hides, the hidden finger posture is estimated using the above method.

By the way, the same finger joint angle may be continued by the finger motion stopping in past frames. In this case, the polynomial showing transition of finger joint angle in past frames becomes a constant. Therefore, only fixed finger joint angle can be estimated. Then, the finger joint angle data in past frames to refer to is not updated when a finger joint angle is the same as previous frame. However, if it was separately updated about each finger joint, lack of data might

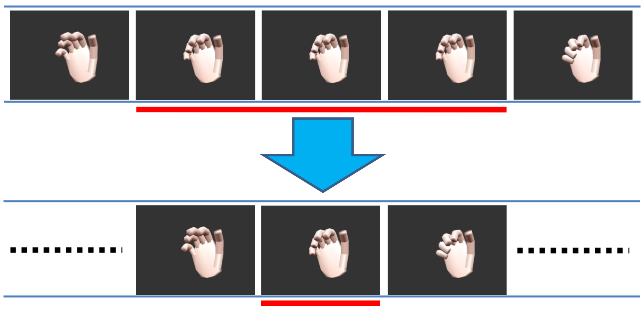


Figure 3: Updating the finger joint angle data in past frames. The same finger joint angle with three of top middle. Leave one of three with bottom middle.

occur when the correlation is derived. The finger joint angle data in correlation is not updated only when both of finger joint angles do not change. When data is not updated, time between frames which is not updated is excepted as shown in Figure 3.

Estimated hidden finger posture has to continue finger motion before hiding. However, since the hidden fingertip is undetectable in fact, the finger motion is not necessarily correct. Then, the hidden finger motion is estimated from estimated hidden finger posture with the polynomial and previous frame posture.

B. Shade Space of Self-occlusion

The hidden finger motion is estimated using current frame fingertip being hidden with self-occlusion. A fingertip is hidden with self-occlusion. In other words, a hidden fingertip is behind other fingers from a camera viewpoint. There we call the space behind other fingers as shade space. Collision detection of shade space and fingertip position of previous frame posture is performed. And collision detection of shade space and estimated hidden fingertip position with the polynomial is performed. The hidden finger motion is estimated from the results. However, shade space is complicated-shaped 3D space. Therefore computational complexity is large when collision detection is performed as 3D space.

When the hidden fingertip is in shade space, a line segment which connects the hidden fingertip and a camera crosses the finger which makes shade space. We thought that a finger is cylinder and a finger joint is sphere joint. A finger consists of phalanges. A phalanx can be expressed with a capsule as shown in Figure 4. Thus, collision detection is possible using the distance calculation between a line segment and axis of phalanx (Figure 5.).

When $D \leq r$, L_s and L_A cross and the fingertip is in shade space. By performing the distance calculation to all shade space, it is judged whether the hidden fingertip is in shade space.

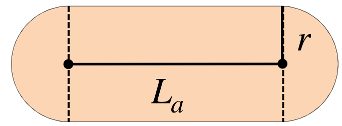


Figure 4: Phalanx. r is radius. L_a is axis of the cylinder.

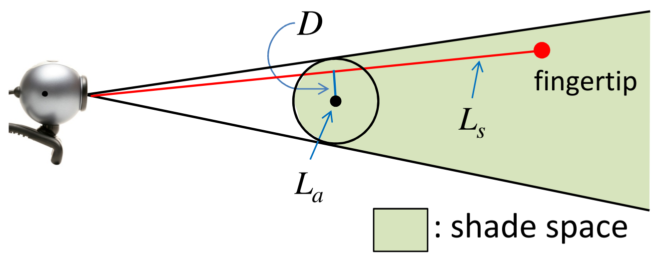


Figure 5: Collision detection. L_s is a line segment which connects the fingertip and the camera. D is the distance between L_s and L_a.

C. Hidden Finger Posture Estimation Considering Finger Motion

The hidden finger motion is judged as follows from the results of collision detection:

- Finger motion which completely continued the motion before the fingertip hides.
- Finger motion slower than the motion speed before the fingertip hides.
- Finger motion different from motion before the fingertip hides.
- Finger motion is stopped.

The hidden finger posture is estimated from this judgment.

1) Two Fingertips are Inside: When two fingertip positions are in shade space, the hidden finger motion is moving within shade space or is stopped.

First, we assume that the hidden finger is moving within shade space. By the way the hidden finger posture is estimated based on transition of the finger joint angle which is showed with the cubic polynomial. When the hidden finger is in shade space over several frames, the hidden finger moves along with the cubic polynomial. When the polynomial is drawn like sine curve or cosine curve, the hidden finger motion is repetitive motion within shade space as shown in Figure 6.

However, when the hidden finger is in shade space over several frames, it is rare that the hidden finger moves repeatedly as mentioned above. In such a case, other fingers are moving. Since shade space is moved, the hidden finger moves along with motion of shade space, without repetitive motion. Or the fingertip comes out of shade space and is detected. Thus, the hidden finger does not move repeatedly in shade space. This is to say movement direction of the hidden finger is not reversed.

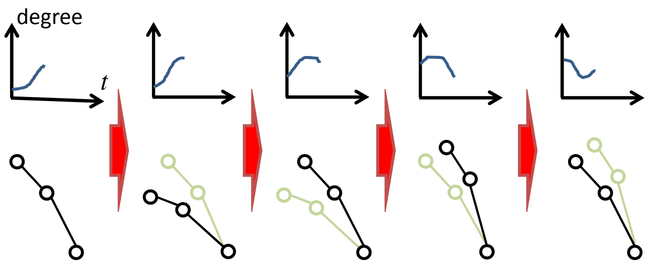


Figure 6: Repetitive finger motion.

When estimated finger posture with the polynomial is the finger posture makes the movement direction reverse, it is thought that the hidden finger posture is target finger posture. It would be better to say that the hidden finger motion stopped. In such a case, let previous frame posture be hidden finger posture in current frame. When estimated finger posture with the polynomial is finger posture in which the motion before the fingertip hides is continued, estimated finger posture with the polynomial is the hidden finger posture in current frame.

2) Estimated Fingertip is Inside, Previous Frame Fingertip is Outside: When estimated hidden fingertip position with the polynomial is in shade space and fingertip position of previous frame posture is out of shade space, the hidden finger is changed into previous frame posture from the estimated finger posture with the polynomial. Thus, estimated finger posture with the polynomial is the hidden finger posture in current frame.

3) Estimated Fingertip is Outside, Previous Frame Fingertip is Inside: When estimated hidden fingertip position with the polynomial is out of shade space and fingertip position of previous frame posture is in shade space, the hidden finger motion is slower later than the motion speed before the fingertip hides, is stopped or that movement direction is reversed. Since it becomes repetitive motion when movement direction is reversed, it is excepted. When the finger motion speed is 0, the finger motion is stopped. Thus, the hidden finger posture can be estimated with the cubic polynomial.

The hidden finger posture estimation with the polynomial is given current time t_n to the polynomial. By changing time given to the polynomial, the hidden finger posture of motion to which speed becomes slow can be estimated. The time given to the polynomial is from current time t_n to previous time t_{n-1} for every short time δ. Let a finger posture with the fingertip position which near most current time is judged to be inside of shade space be hidden finger posture in current frame.

4) Two Fingertips are Outside: When two fingertip positions are out of shade space, the movement direction of the hidden finger is reversed or the finger motion is changed to different motion.

First, we thought that case where the movement direction is reversed. In this case, since fingertip position of previous frame posture is outside shade space, it is permissible that the movement direction of the hidden finger is reversed. The hidden finger posture in this case is estimated using the same method as the case where speed becomes slow. The time given to the polynomial is $t_{n-1} - \Delta t$. Time Δt is a time from t_{n-1} to t_n. The hidden finger posture in the finger motion to reversal direction of the finger motion is estimated with Δt. Collision detection of shade space and the fingertip position of estimated finger posture with Δt is performed. When the result is in shade space, the hidden finger moves in the reversal direction. Let estimated finger posture with Δt be hidden finger posture in current frame.

When the fingertip position of estimated finger posture with Δt is outside shade space, the finger posture does not satisfy a condition in which the fingertip hides with self-occlusion. Thus, It is thought that hidden finger motion was changed. The hidden finger posture in this case is estimated based on image coordinates from fingertip position of previous frame posture and shade space.

D. Hidden Finger Posture Estimation of Other Motions

When changed to motion from which the hidden finger motion is different, it is difficult to estimate the hidden finger posture from time series data. Therefore, the hidden finger posture is estimated based on image coordinates from the hidden fingertip position of previous frame posture and shade space. In order to show shade space correctly using image coordinates, we have to project form of phalanx of all detected fingers. However, computational complexity will increase. For computational complexity reduction, shade space in image coordinates is shown with a point P on which the central point of phalanx which builds shade space is projected. A phalanx which makes shade space where the hidden fingertip exists is estimated based on P and the fingertip position of previous frame posture. The hidden finger posture is estimated based on estimated phalanx.

Point Q is the point which is projected the fingertip position of previous frame posture on image coordinates. Distance between P and Q on image coordinates is calculated. The hidden finger posture is estimated assuming that the fingertip exists in shade space which phalanx of P with shortest distance makes. Distance between a line which connected a camera and the central point of phalanx, and the fingertip position of of previous frame posture is calculated. Let the point on the line which becomes shortest distance be fingertip position in current frame. the hidden finger posture is estimated with solving IK from a fingertip position.

When the distance between the point on the line and the fingertip position of solution to IK is below a threshold value, let the fingertip position of solution to IK be hidden finger posture in current frame. When the distance is larger than the threshold value, the fingertip does not reach

Table II: Estimation Order

Thumb	Front side	Order
detect	Index	Index → Middle → Ring → Little
detect	Little	Little → Ring → Middle → Index
undetect	Index	Index → Middle → Ring → Little → Thumb
undetect	Little	Little → Ring → Middle → Index → Thumb

Table III: Process Time

Algorithm	Time (msec)
Camera image fetch and wait	16.2
Noise reduction	13.7
Fingertip position calculation	0.6
Hand posture estimation	2.5
CG drawing	0.3
Total	33.3

the point on the line with hand structural constraint. The hidden finger posture is estimated with phalanx with next short distance on image coordinates. When the distance in all phalanges becomes more than the threshold value, we thought that the fingertip is out of a image. Thus, let previous frame posture be finger posture in current frame.

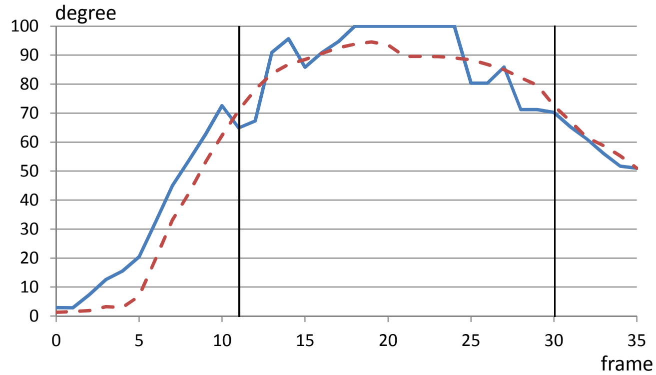
E. Several Hidden Finger Posture Estimation

When there are some hidden fingers, a hidden fingertip may exist in shade space of other hidden finger. A finger posture must be estimated from fingers in front side. As shown in Table II, an order of finger posture estimation is determined from direction of the hand and hand structural constraints.

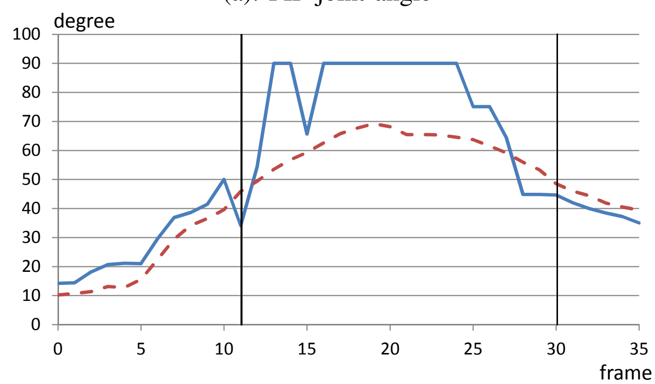
IV. Experimental Result

A experimental system using the method mentioned above is implemented with C++ language. The specification of PC is the following; CPU: Dual-Core AMD Opteron Processor 1210 1.8GHz. The camera is BSW20K04 of BUFFALO INC. (MAX fps: 30 fps). We evaluated process speed and estimation accuracy of our experimental system. The results of process time are shown in Table III. The average process speed was the same 30 fps as the camera operation maximum. The camera image fetch and wait time was 48% of them and the noise reduction of an image was 41%. The speed is sufficient and it is useful as an interface.

To evaluate the estimation accuracy, it estimates the same hand posture from two camera images. All of fingertips are taken in one image, and the estimated value from this image is assumed as true. Some of fingertips are hidden on the other image, then the estimated posture is compared with the true one. Both of cameras are set at a distance of 1m from the hand, and the angle between the camera directions is 40 degrees. The results of estimation accuracy and result images are shown in Figure 7,8.



(a): PIP joint angle



(b): MP joint angle

Figure 7: The results of estimation accuracy. A hidden finger is index finger. Index finger is hidden (11–30frames). A dashed line is true value and a solid line is estimation value.

When all fingertips are detected from two viewpoints, the mean errors are 6.77 degrees at PIP joint angle, and 5.60 degrees at MP joint angle. When hidden finger posture is estimated, the mean errors are 6.78 degrees at PIP joint angle, and 18.10 degrees at MP joint angle. The estimation of our system is sufficient accuracy. However, the estimate error of MP joint angle is slightly large. The estimated value was more bent than the true value. Although it is not bent to maximum of the movable range in fact, the system probably estimated like that the finger motion was still continued and it bent to maximum of the range. We expect that suitable and flexible movable range considering other postures of visible fingers works better.

V. Conclusion

In this paper, we proposed a new Vision Based Data Glove considering hidden fingertip with self-occlusion. It can estimate hand posture (finger joint angles) on the assumption that fingertips are already detected using computer vision technology, even when a camera cannot detect some fingertips with self-occlusion.

The experiment result shows that the system can estimate a finger posture with self-occlusion based on a finger

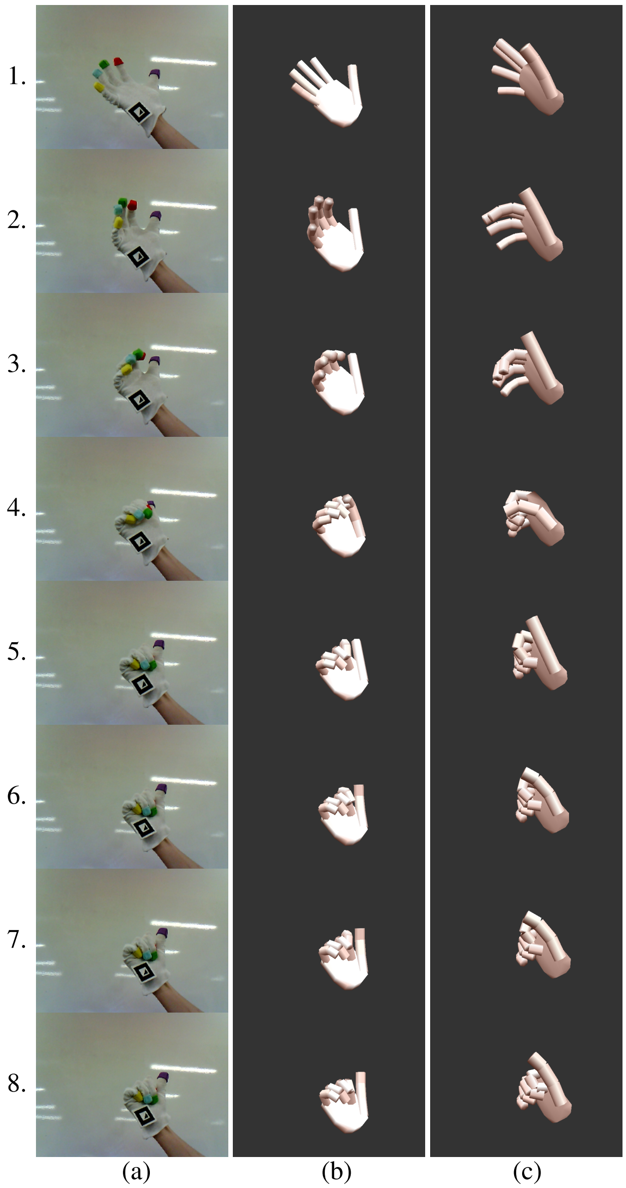


Figure 8: The result of estimation. The figure (a) is camera images. The figure (b) is the result of estimation drawn in CG from the camera viewpoint. The figure (c) is the result of estimation drawn in CG from another viewpoint. Index finger is detectable at 1–4. Index finger is undetectable at 5–8. However index finger posture can be estimated at 5–8.

motion with sufficiently high process speed. Although our experiment system detects fingertips using some markers, the estimation method is applicable to other types of VBDG system which detect fingertips based on other techniques.

In the future, we would like to estimate hidden finger posture not only the finger motion but also other fingers and relations with them. We will also apply the fingertip detection method without any markers to our system. It is expected that VBDG spreads through home as an inexpensive interface for VR system.

REFERENCES

[1] Emi Tamaki, Takashi Miyata and Jun Rekimoto: A Robust and Accurate 3D Hand Posture Estimation Method for Interactive Systems, Transactions of Information Processing Society of Japan (in Japanese), Vol. 5, No. 2, pp. 229-239, 2010.

[2] Robert Y. Wang and Jovan Popovic: Real-Time Hand-Tracking with a Color Glove, ACM Transaction on Graphics (SIGGRAPH 2009), Vol. 28, No. 3, 2009.

[3] B. Stenger, A. Thayananthan, P. H. S. Torr, R.Cipolla: Model-Based Hand Tracking Using a Hi-erarchical Bayesian Filter, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No. 9, pp. 1372-1384, 2006.

[4] M.K. Bhuyan, Debanga Raj Neog and Mithun Kumar Kar: Hand Pose Recognition using Geometric Features, National Conference on Communications (NCC), 2011.

[5] Kenji Oka, Yoichi Sato, and Hideki Koike: Real-time fingertip tracking and gesture recognition, IEEE Computer Graphics and Applications, Vol. 22, No. 6, pp. 64-71, 2002.

[6] Vitor F. Pamplona, Leandro A. F. Fernandes, João Prauchner, Luciana P. Nedel e Manuel M. Oliveira: The Image-Based Data Glove, Proceedings of X Symposium on Virtual Reality (SVR'2008), João Pessoa, 2008. Anais do SVR 2008, Porto Alegre: SBC, 2008, (ISBN: 857669174-4). pp. 204-211.

[7] Weiying Chen, Ryuji Fujiki, Daisaku Arita and Rin-ichiro Taniguchi: Real-time 3D Hand Shape Estimation using Multiple Cameras, Proceedings of the 13th Japan-Korea Joint Workshop on Frontiers of Computer Vision, pp. 15-20, 2007.

[8] Ryuji Fujiki, Daisaku Arita, Rin-ichiro Taniguchi: Real-time 3D hand shape estimation based on inverse kinematics and physical constraints Proceedings of International Conference on Image Analysis and Processing, pp. 850-858, 2005.

[9] Sanshiro Yamamoto and Kenji Funahashi: A Study for Image Based Data Glove to Use at Personal Home, Proceedings of the 15th Annual Conference of the Virtual Reality Society Japan, 2A2-5 (DVD-ROM, in Japanese), 2010.

[10] Hirokazu and Mark Billinghurst: Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System, Proceedings of the 2nd International Workshop on Augmented Reality (IWAR99), pp. 85-94, 1999.

[11] Chris Welman: Inverse kinematics and geometric constraints for articulated figure manipulation, M.Sc Thesis, Simon Fraser University, 1993.