Indication of Object Spatial Position by Finger Pointing

Masato HIRASAWA
Graduate of Mercantile Marine
Tokyo University of
Marine Science and Technology

Masaki OSHIMA
Faculty of Marine Technology
Tokyo University of
Marine Science and Technology

Tokyo-to Koto-ku Etchujima 2-1-6 〒135-8533

***Key Words*: Spatial Coordinate, Finger Pointing, Conversation, Indication**

Abstract

Human often communicate with each other by gesture for natural communication. The target of this study is to provide a technology that enables to communicate with robot by finger indication in natural way.

In order to set the direction of finger to one, at least two spatial coordinates are required on the finger. In this research, spatial coordinate is acquired using stereo method. This spatial straight line is called “Finger indication vector”.

The finger area is extracted by skin color and shape feature of hand. From the area of the extracted finger, the portion of the tip of a finger and the root of a finger are made into the corresponding points.

Plane information is calculated by normal vector. And the object is pointed out by connecting intersection of finger indicated vector and object plane, and finger.

The result of experiment is shown finger indication vector run out in the portion without feature in object plane. It is shown in 3-dimensional space that the finger indication vector and object plane crossed. Even if this technique does not put on special equipment, it can presume the direction of finger, and subject is able to experiment it in free state.

1. Introduction

In order to support life and work of human, the robot that coexists with people have developed [1],[2]. Human must understand with each other for coexistence in everyday life. Human often communicate by “Hearing”, ”Looking”, ”Speaking”. Study for human and robot talking communication have been done by language such as character and sound [3][4]. As oppose to this, the technology to communicate with robots have been developed by non-language such as gesture [5][6]. In order to communicate in natural, human often communicate with each other by gesture. For example, when speaker wants to take the object in the position which speaker left, he says “Please take that object to me” while it points out [7]. The system which holds a conference as if people who are present in a mutually different place meet with a same place is proposed. One of the factor reproducing presence, it is made important that enables to point out by finger through a screen. In order to communicate in natural, this also shows that finger pointing is important. It relates to this, the research which teaches objective fields and the position of the peak using the laser pointer [8], and the research which detects the gazing point from a look [9] have accomplished.

Even if it dose not use special device, if it is human, it is possible by directing the target object with finger to aim at communication. The target of this research is to provide a technology that enables to communicate with robot by finger pointing in natural way. In order to know spatial position between finger and object, hand and object are horizontally taken using stereo camera. Spatial coordinate of finger and object are acquired based on stereo photos and indication direction is appointed. And, in order to confirm validity, the experiment that directed object is made to answer computer is conducted.

2. Decision of Finger Indication Vector

2.1 Finger Indication Vector

Indication direction is conducted by attached spatial straight line on finger. In order to set the direction of finger to one, at least two spatial coordinates are required on the finger. In this research, spatial coordinate is acquired by stereo method. The direction of finger is presumed by applying straight line on these spatial coordinates. Straight line vector is calculated by formula (1). This is called “Finger Indication Vector” in this research.

$$\vec{r} = \vec{b} + t \cdot \vec{a} \qquad (1)$$

However, $\vec{a}$ and $\vec{b}$ are expressed “Finger Indication Vector”, and t is expressed real number.

2.2 Acquisition of hand area

After moving object is extracted by inter-frame difference, each extracted areas are done labeling, and small areas are removed as noise. Skin color area is extracted from left-behind area and it is considered that recognized as hand area. Inter-frame difference tends to be influenced of noise, even if indoor environment. It is caused incorrect detection by sunlight and fluorescent light. If it thinks that density pattern of picture does not change, incorrect detection can be reduced by making luminosity value normalization like formula (2).

$$f_i'(x,y) = \frac{f_i(x,y)}{\sqrt{\sum_{x,y}\{\overline{f_i(x,y)} - f_i(x,y)\}^2}} * k \qquad (2)$$

However $f(x,y)$ is expressed luminosity, $\overline{f(x,y)}$ is expressed average, and k is expressed real number.

Color information is changed Hue by formula (3), and skin color area is extracted. However R,G,B and $C_{\max}, C_{\min}$ are expressed color information and value of maximum or minimum into R,G,B, respectively, and H is expressed value of Hue.

$$H = \begin{cases} if \quad G = C_{\max} \\ (B-R)/(C_{\max} - C_{\min}) * 60 + 120 \end{cases}$$
$$H = \begin{cases} if \quad B = C_{\max} \\ (R-G)/(C_{\max} - C_{\min}) * 60 + 120 \end{cases} \qquad (3)$$
$$H = \begin{cases} if \quad R = C_{\max} \quad B > G \\ (G-B)/(C_{\max} - C_{\min}) * 60 + 360 \end{cases}$$
$$H = \begin{cases} if \quad R = C_{\max} \quad G > B \\ (G-B)/(C_{\max} - C_{\min}) * 60 \end{cases}$$

2.3 Determination of finger direction

The flow chart of algorithm, which determines finger direction, is shown in Fig.2. The direction of hand (horizontal or vertical) is determined by form of the rectangle of the extracted hand area. If it is in the tendency for the hand to be horizontally suitable, it scans vertically, and the number of pixels of the pixel value 255 is recorded an account for the every sequence. By the middle of scan, the boundary line of finger and back of hand is made the portion which count value changed a lot. If the hand is vertically suitable, the scanning direction is changed and it processes similarly. If difference of vertically direction and horizontally direction are not clear, it corresponds scanning aslant.

2.4 Determination of spatial coordinate on finger

Fig.1 The shape of hand

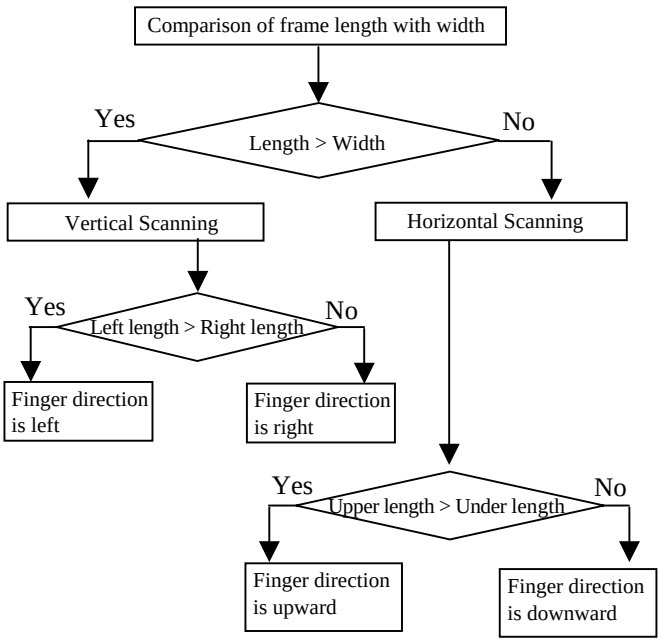

Fig.2 The flowchart of algorism

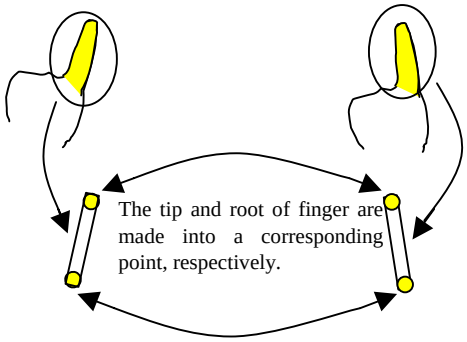

Fig.3 Extraction of correspondence points

In order to determinate direction of finger indication vector, two spatial coordinates are search on the finger. Since it calculates by stereo method, corresponding points have to be searched in the picture on either side. From the area of the extracted finger, the portion of the tip of a finger and the root of a finger are made into the corresponding points, respectively, as depicted by Fig.3. Formula (1) is calculated by searched two spatial coordinates.

Cameras are used the lens of the same characteristic and arranged it in the same height. Starting point of

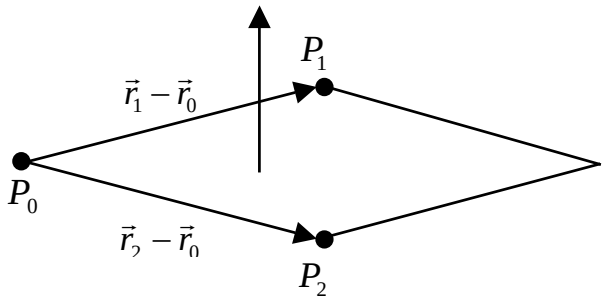

Fig.4 The Normal

spatial coordinate is taken as the position of lens on the left-hand side of a stereo camera.

3. Detection of object position

By the experiment, which searches intersection with an object plane, it shows validity of finger indication vector searched section 2. Although the plane information is able to search for also by method by reference [10], this paper presumes plane information by searching for the three spatial coordinates, which exists on the same plane (Fig.4). First, plane equation is expressed formula (4). ρ_0 is expressed normal and can calculate it like formula (5). Normal ρ is able to calculate by substituting one of value $\vec{r}_1, \vec{r}_2$ or $\vec{r}_3$. Here it is necessary to search for three spatial coordinates, which exist on the same plane. Then three feature points, such as objective edges, are extracted on the same plane. Next, three spatial points are acquired by taking correspondence between the pictures obtained from the camera on either side. Formula (6) is drawn from formula (1) and (4), coefficient can be calculated. Intersection $\vec{r}$ can be obtained by substituting for formula (1) the value calculated here. The object is pointed out by connection finger indicated vector and intersection $\vec{r}$.

$$\vec{r} \cdot \rho_0 = \rho \tag{4}$$

$$\rho_0 = \frac{(\vec{r}_2 - \vec{r}_0) \times (\vec{r}_1 - \vec{r}_0)}{\left|(\vec{r}_2 - \vec{r}_0) \times (\vec{r}_1 - \vec{r}_0)\right|} \tag{5}$$

$$t = \frac{\rho - \rho_0 \cdot \vec{b}}{\rho_0 \cdot \vec{a}} \tag{6}$$

4. The experiment

4.1 The environment of experiment

In order to discover object, the experiment about finger pointing is conclude. The picture photo by the stereo camera (3DC-2000Z SONY) is send to computer through a picture edit machine (Accom-WSD-2Xtreem NGC) , and result is outputted. An experiment is conducted so that the hand and object are taken under complicated background. Objective spatial position is acquired beforehand and memorized. The target object is premised on not changing position frequency as if home electronics. But, if a position changes, moving area is extracted and memory is updated.

4.2 Evaluation of finger pointing

The experiment for confirming the accuracy of a finger indication vector is conducted by comparing actual spatial angle with the direction of vector calculated from reproduced spatial coordinates. In the experiment, a laser pointer is fixed to a tripod and a gazing board is installed in the position 1m away from the tripod. The actual angle is determined by attaching gazing points to the interval of 5 degrees vertically, attaching to the interval of 10 agrees horizontally, and irradiating each gazing points by laser point. This state is photoed and direction vector is calculated by formula (7) and (8). Here, (X_1, Y_1, Z_1), (X_2, Y_2, Z_2) expresses there produced spatial coordinates. The graph which compared actual measurements with calculation values is shown in Fig.5.6. An actual measurement is shown in vertical axis and a calculation value is shown in horizontal axis. The graph expresses comparison with the straight line of ideal which the error of angle is 0 °, and the straight line obtained by experiment. Moreover, the error of finger pointing of the 1[m] beyond, 5[m] beyond, and each 10[m] beyond which the error of angle of an angle brings about, is shown in Table1 and 2.

$$\vartheta = \tan^{-1}\left(\frac{X_2 - X_1}{Z_2 - Z_1}\right) \tag{7}$$ （Vertically）

$$\vartheta = \tan^{-1}\left(\frac{Y_2 - Y_1}{Z_2 - Z_1}\right) \tag{8}$$ （Horizontal）

4.3 The result of finger pointing

The result of experiment is shown in the Fig.7 and 8. Each experiment shows the result by projecting the spatial vector obtained by the formula (1) on plane coordinate. After extracting also for an intersection spatially, it is produced on plane coordinates. In this figure, finger indication vector run out in the portion without feature in object plane. It is shown in 3-dimensional space that the finger indication vector and object plane crossed. The desirable method for state that

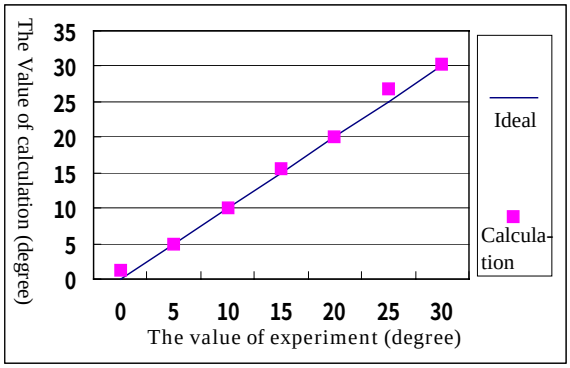

Fig.5 Calculation value and actual measurement of vertical

Table.1 Digital data of Fig.6

Actual measurement (degree)	Calculation value (degree)	Angle error (degree)	The point error of 1m beyond (mm)	The point error of 5m beyond (mm)
0	1.23	1.23	21.54	107.70
5	4.97	-0.03	-0.52	-2.62
10	9.98	-0.02	-0.38	-1.92
15	15.61	0.61	10.68	53.41
20	20.07	0.07	1.19	5.93
25	26.89	1.89	33.07	165.34
30	30.39	0.39	6.86	34.30

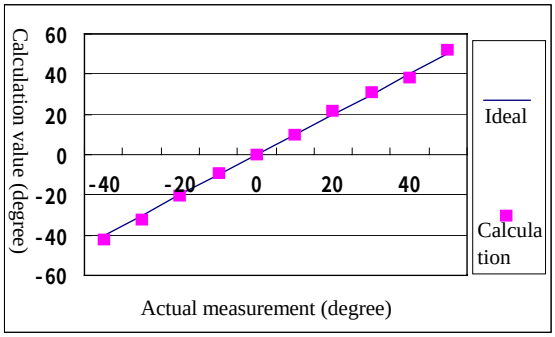

Fig.6 Calculation value and actual measurement of horizontal

Table.2 Digital data of Fig.7

Actual measurement (degree)	Calculation value (degree)	Angle error (degree)	The point error Of beyond (mm)	5m 先のポイント誤差 (mm)
-40	-42.34	-2.34	-40.86	-204.32
-30	-32.42	-2.42	-42.25	-211.22
-20	-20.21	-0.21	-3.68	-18.41
-10	-9.26	0.74	12.90	64.49
0	-0.06	-0.06	-1.01	-5.062
10	9.70	-0.31	-5.32	-26.62
20	21.89	1.89	32.91	164.56
30	31.28	1.28	22.29	111.46
40	38.53	-1.47	-25.68	-128.40
50	52.03	2.03	35.46	177.31

Fig.7 Example of pointing (TV)

Fig.8 Example of pointing (Display)

object is directed being shown is giving definition as a polyhedron using a solid modeler, and judging the interaction with the finger indication vector and object plane. Here, since it is simple, only a certain field on object is taught, only the spread range is defined, and it is considered that the object is directed only within the intersection within the limit of it.

5. Conclusion

In this report, finger indication is recognized from the picture obtained by the stereo camera, and object is directed. Even if this technique does not put on special equipment, it can presume the direction of finger, and subject is able to experiment it in free state. Position error and angle error is a maximum of 211 [mm] and 2.42 degree, respectively in the distance of 5[m] around. As a result, if it is a size like home electronics, the objective position is able to be taught. It is necessary to raise pointing accuracy so that it can presume, even if it is the small object that distance left from now on. Whether the finger is correctly suitable in the direction of objective influences evaluation greatly. As a subject, the rough instruction of “being around here generally” is also construction of a system by which information is transmitted.

Reference

[1] T.Takahashi, S.Nakanishi, Y.Kuno and Y.Shirai：“ Helping

Computer Vision by Verbal and Nonverbal Communication", International Conference on Pattern Recognition, Vol.2, pp.1216-1218, 1998

[2] C.Breazeal and B.Scassellati: "How to build robots that make friends and influence people", Proc. IEEE/RSJ International. Conference. Intelligent Robots and Systems, 1999

[3] K.Shobaki, J.P.Hosom and R.A.Cole: "The OGI Kid's speech corpus and recognize" Proc. ICSLP2000, Vol.4, pp.258-261, 2000

[4] K.Ohtsuki, T.Matsuoka, S.Matsunaga and S.Furui:"Topic extraction based on continuous speech recognition in broadcast news speech", IEICE Trans. Inf. & Syst., Vol.E85-D, no.7, pp1145-1152, July, 2002

[5] V.I.Pavlovic, R.Sharma and T.S.Huang: "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review", *IEEE Trans.Pattern Analysis and Machine Intelligence*, Vol.19, No.7, pp.677-695, 1997, 7

[6] A.Pentland: "Looking at People: Sensing for Ubiquitous and Wearable Computing", *IEEE Trans.Pattern Analysis and Machine Intelligence*, Vol.22, No.1, pp.107-119, 2000, 1

[7] O. Morikawa and T. Maesako: "HyperMirror :Toward Pleasant-to-use Video Mediated Communication System", CSCW'98, pp.149-158, 1998

[8] T.Hasegawa and S.Kameyama: "Geometric Modeling of Manipulation Environment with Interactive Teaching and Automated Accuracy Improvement", 20th Int. Symp. on Industrial Robots, pp.419-426, 1989

[9] Jacob. R: "What you look at is what you get: Eye movement-based interaction techniques", Proc. ACM CHI'90 Human Factors in Computing System Conference, pp.11-18, 1990

[10] Y. Kawai, Y. Ueshiba, T. Yoshimi and M. Oshima: "Reconstruction of 3D Objects by Integration of Multiple Range Data", ICPR(International Conference on Pattern Recognition), Hague, The Netherlands, pp.154-157, 1982, 9