HOW TO SELECT USEFUL HAND SHAPES FOR HAND GESTURE RECOGNITION SYSTEM

Atsushi Shimada[1], Takayoshi Yamashita[2] and Rin-ichiro Taniguchi[1]
[1]*Graduate School and Faculty of Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan*
[2]*OMRON Corporation, Shiga, Japan*

Keywords: Hand gesture recognition, Selection, Supporting system.

Abstract: This paper discusses hand shapes for Human Computer Interface. Usually, a hand gesture based Human Computer Interface is developed by human centered design concept. A system designer or developer tends to select hand shapes by himself/herself without verifying practical effectiveness from the standpoint of system aspect. Instead, a methodology of training and recognition of hand shapes is often discussed. On the other hand, this paper listens to system's voice; which hand shape is easy to be recognized, which is easy to be confused and so on. Actually, 37 kinds of tentative hand shapes were investigated from the viewpoint of system-friendly hand shape. Based on the result, a supporting system was developed for a system designer, which helps to find appropriate hand shapes which satisfy both "user-friendly" and "system-friendly" demand.

1 INTRODUCTION

Recent years, hand gesture has been widely used for Human Computer Interface(HCI). Vision-based approach is an attractive way to realize hand gesture applications since people can send a command to a computer without putting a special sensor (such as hand glove, physical sensor and so on) on his/her hands (Zabulis et al., 2009; Wachs et al., 2011). Generally, vision-based hand gesture recognition can be divided into two issues; a motion issue (Chen et al., 2008) and a shape issue (Wang and Wang, 2007). In some literatures, these issues are discussed in the lump. This paper is concerned with a hand shape issue.

There are mainly two approaches to recognize a hand shape. One is a model-based approach (de La Gorce et al., 2008) and the other is an appearance-based approach (Martin and Crowley, 1997). This paper focuses on the appearance-based hand shape recognition. One of the straight forward ways is to let a system train a lot of training samples including various changes of appearance. When the number of shapes is relatively small, such an approach works well. However, if many kinds of shapes are required, some shapes are sometimes confused as other shapes. The problem is often caused by the rotation of the wrist. When the wrist rotates to a certain direction, the appearance will change, and the changed appearance sometimes becomes similar with other hand shape. Traditional researches tended to select expedient shapes and/or set a constraint on the rotation to avoid this problem.

This paper discusses a matter of selecting hand shapes, which is absolutely imperative to develop a HCI based on hand gesture. Although many literatures discuss the novel methodology of recognition, extraction of effective/powerful visual features and so on, this paper does not treat such a theme. Instead, this paper backs to basics of system design and provide a discussion about what kinds of hand shapes are useful for hand-gesture based applications.

When a new hand-gesture based system is designed, two aspects should be considered. If a user or a system designer can liberally choose hand shapes which he/she wants to use, it is a favorable condition from a human stand point. We call them "user-friendly" hand shapes. On the other hand, a system would hope discriminative hand shapes, which are called as "system-friendly" hand shapes in this paper. Therefore, a system designer has to search for common ground which satisfies both requirements. To help such a system designer, this research provides a supporting system for hand shape selection. All the designer has to do is to select some hand shapes as he/she like. The system feeds back five barometers based on the selected hand shapes. Through the interaction with the system, a designer can find a meeting point, which means that the system accepts the

Shimada A., Yamashita T. and Tanguchi R. (2012).
HOW TO SELECT USEFUL HAND SHAPES FOR HAND GESTURE RECOGNITION SYSTEM.
In *Proceedings of the 1st International Conference on Pattern Recognition Applications and Methods*, pages 394-399
DOI: 10.5220/0003767603940399

Figure 1: Environment of capturing hand shapes

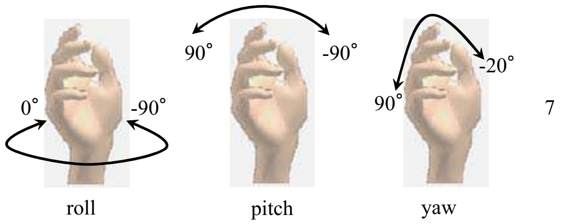

Figure 2: Rotation ranges

request from the designer as much as possible and the designer also accepts the system's requirement.

2 EVALUATIVE ENVIRONMENT

To realize a supporting system for hand shape selection, 37 kinds of hand shapes with various rotations were captured by a camera. Then, visual similarity within/among shapes was evaluated. Besides, recognition accuracy was investigated. These evaluations were conducted under three situations.

Case 1 all shapes with all rotations

Case 2 limited shapes with all rotations

Case 3 all shapes with limited rotations

2.1 Rotation Ranges of Wrist

As shown in Figure 1, hand shapes were captured by a camera fixed in front of the hand. The arm was fixed by handmade equipment. The appearance of hand shapes could be changed by three rotations as roll, pitch and yaw(see Figure 2) constrained by the ranges below.

roll: clockwise rotation, $-90^\circ \sim 70^\circ$

pitch: front-back direction, $-90^\circ \sim 90^\circ$

yaw: horizontal direction, $-20^\circ \sim 90^\circ$

Note that these rotation ranges cover the range of movement of human wrist. Some examples are shown in Figure 3.

2.2 Hand Shapes

Our research picked up 37 hand shape classes(see Figure 4) which are used in sign language communication in Japan. The selected hand shapes can be classified into three groups; 23 kinds of shapes classes in which fingers point upward, sideways for 7 shape classes and downward for 7 shape classes. As we can see, there are some similar shapes even though all shapes are captured in basic positions (i.e. no rotation) in Figure. 4. Besides, some hand shapes become similar if the shapes rotate in the yaw direction (e.g. class 25 and class 34).

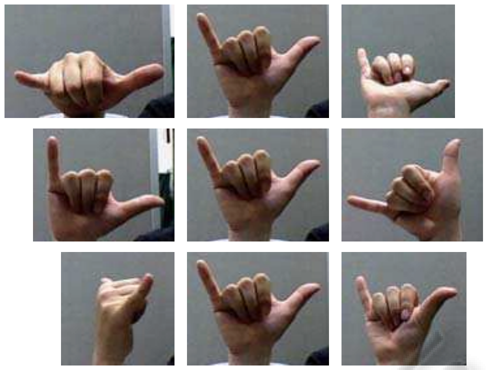

Figure 3: Examples of captured hand shapes

2.3 Detection of Hand Shape Region

Hand shape masks were detected from about 20,000 images collected by 4 examinees. The captured image size was 640x480, and the detected region was resized into 100x100 pixels. Then, the image was converted to gray scale followed by histogram normalization.

2.4 Evaluation of Similarity

The similarity was evaluated by L2 norm between two hand shape images, i.e. accumulation of L2 distances between corresponding pixels. To acquire the similarity score within the class, the average L2 norm was calculated from all possible combinations. The smaller the average L2 norm is, the less the appearance changes within the class. On the other hand, to acquire the similarity score between two classes, the L2 norm was calculated between representative shapes of each class. The representative shape was acquired by averaging all images within the class. To avoid the conflict among shape classes, it is better for the score to be large. Based on these two scores, the conflict score is defined by

$$E_{i,j} = \frac{d_i + d_j}{D_{i,j}} \tag{1}$$

where d_i and $D_{i,j}$ are the L2 norm within the class i and between the class i and j respectively. With de-

Figure 4: 37 Kinds of hand shapes

creasing the $E_{i,j}$, corresponding two classes i and j can be distinguished well from each other. Besides, the class score is defined by

$$S_i = \frac{1}{N}\sum_j E_{i,j}. \tag{2}$$

The smaller score S_i is preferable from the viewpoint of voiding the confusion with other shape classes.

2.5 Evaluation of Recognition Accuracy

Recognition accuracy is investigated through leave-one-out cross validation. The recognition is performed by searching the most similar image in the training samples. To avoid the curse of dimensionality, all the images are analyzed by PCA(Principal Component Analysis) and the number of dimensions was reduced from 10,000 (100x100) to 100 before recognition process.

3 EVALUATION RESULTS

3.1 Experimental Procedure

As described in section 1, three situations are assumed in the experiment. Here, the more detailed conditions are described as follows.

Case 1 all 37 hand shape classes with all rotations

Case 2 selected 20 hand shape classes with all rotations

Case 3 all 37 hand shape classes with limited rotations

In each case, the conflict score $E_{i,j}$ and class score S_i are calculated.

In the Case 1, all shapes with all rotations are evaluated. Based on the result of Case 1, 20 hand shape classes are selected in the Case 2. In the Case 3, the rotations of the wrist are restricted as follows.

roll: clockwise rotation, $-20° \sim 20°$

pitch: front-back direction, $-20° \sim 20°$

yaw: horizontal direction, $-20° \sim 20°$

Examples of hand shapes in the condition of Case 3 are shown in Figure 5. The appearance change in Figure 5 becomes smaller compared with the images shown in Figure 3.

3.2 Results

The class score S_i and recognition accuracy are shown in Table 1 and Table 2. In the Case 2, 17 shape classes were excluded from evaluation. In Table 1 and Table 2, the scores of excluded classes are drawn by the "–" mark. For example, the class 4 and 5 had a

Table 1: Class score.

Class	Case 1	Case 2	Case 3	Class	Case 1	Case 2	Case 3
1	5.3	–	–	20	3.6	3.6	2.1
2	3.9	3.5	2.4	21	3.3	3.3	2.1
3	3.9	3.5	2.3	22	3.4	3.3	2.0
4	4.8	4.7	3.3	23	4.0	4.1	2.3
5	3.9	3.6	2.3	24	2.7	–	–
6	3.5	3.4	2.3	25	2.3	2.4	1.7
7	3.3	–	–	26	2.1	–	–
8	5.2	–	–	27	2.8	–	–
9	3.8	–	–	28	2.0	–	–
10	3.5	3.4	2.2	29	3.3	2.8	2.0
11	4.0	–	–	30	2.5	2.6	1.9
12	3.8	3.4	2.3	31	3.1	3.2	2.4
13	3.3	–	–	32	3.2	–	–
14	3.2	3.3	2.0	33	3.1	–	–
15	3.6	–	–	34	3.1	3.1	2.2
16	3.6	3.2	2.1	35	3.3	–	–
17	3.6	–	–	36	3.2	–	–
18	3.1	3.0	1.9	37	3.8	3.8	2.6
19	3.7	–	–				

Table 2: Recognition accuracy.

Class	Case 1	Case 2	Case 3	Class	Case 1	Case 2	Case 3
1	0.40	–	–	20	0.53	0.73	0.91
2	0.19	0.40	0.65	21	0.53	0.58	0.79
3	0.27	0.37	0.50	22	0.44	0.45	0.59
4	0.55	0.64	0.81	23	0.69	0.72	0.94
5	0.57	0.67	0.82	24	0.52	–	–
6	0.49	0.52	0.74	25	0.33	0.60	0.96
7	0.58	–	–	26	0.23	–	–
8	0.32	–	–	27	0.14	–	–
9	0.76	–	–	28	0.43	–	–
10	0.41	0.52	0.56	29	0.39	0.73	0.92
11	0.37	–	–	30	0.63	0.70	0.90
12	0.60	0.65	0.69	31	0.34	0.48	0.57
13	0.42	–	–	32	0.40	–	–
14	0.69	0.72	0.86	33	0.47	–	–
15	0.62	–	–	34	0.47	0.72	0.87
16	0.48	0.53	0.64	35	0.25	–	–
17	0.55	–	–	36	0.33	–	–
18	0.57	0.59	0.69	37	0.36	0.42	0.53
19	0.39	–	–				

tendency of conflict with other classes. The conflict of class 5 was resolved in the Case 3 where the rotations were restricted in smaller ranges.

Totally, the class score S_i became smaller from Case 1 to Case 3. Reducing the number of classes and restricting the ranges of rotation contributed to resolve the conflict among classes. Especially, the range limitation produced larger effect compared with reducing the number of classes. With regard to the recognition accuracy, the scores of class 2, 20, 25, 29 and 34 were improved by reducing the number of classes (compare the accuracy score between Case 1 and Case 2). Besides, the Case 3 showed greater improvement of accuracy compared with Case 2.

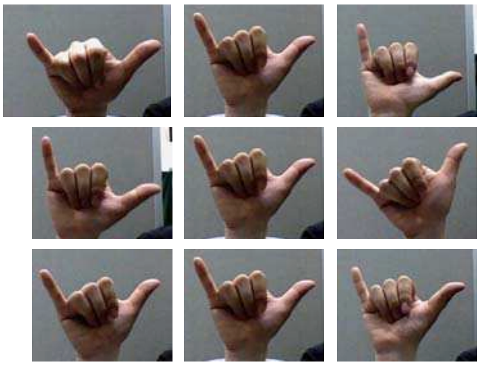

Figure 5: Examples of captured hand shapes with limitation of rotations.

3.3 System-friendly Hand Shapes

Based on the evaluation results above, this paper provides some rankings of "system-friendly" hand shapes. The "system-friendly" score is calculated from the class score S_i and the recognition accuracy P_i as follows.

$$C_i = \frac{P_i}{S_i} \quad (3)$$

The larger value of C_i denotes the hand shape class i will cause less conflict and provide higher recognition accuracy. Therefore, the ranking of system-friendly hand shapes can be made from arranging the C_i in descending order. Table 3 indicates the guide of selecting the hand shapes based on system-friendly ranking.

4 SUPPORTING SYSTEM FOR HAND SHAPE SELECTION

The authors developed a supporting system for hand shape selection[1]. A system designer/ developer can investigate whether the selected hand shapes are system-friendly or not. The user interface of the system is shown in Figure. 6. All the user has to do is to select preferable, i.e. "user-friendly" hand shapes by clicking the hand shape image. The system feeds back five barometers according to the selected hand shapes from the viewpoint of "system-friendly" hand shapes. The barometers consist of following five items.

# of Shapes. The number of shapes selected by the user.

Variation. How wide the rotation ranges are available. (The max score denotes that all rotations are available.)

[1]The system will be available on our Web in the near future.

Table 3: Ranking of System-Friendly Hand Shapes.

Ranking	Case 1	Case 2	Case 3
1	30	30	25
2	28	14	30
3	14	18	29
4	9	23	20
5	24	21	14
6	18	12	23
7	7	34	34
8	23	20	21
9	15	5	18
10	21	25	5
11	12	6	6
12	34	16	16
13	17	22	12
14	33	29	22
15	20	10	2

Separation. How correctly the shape is distinguished from others. (This barometer is calculated based on S_i and $E_{i,j}$.)

Accuracy. How correctly the shape is recognized. (This barometer is calculated by recognition accuracy.)

Total. The comprehensive score based on above four items.

These barometers are drawn on the radar chart(see the bottom-left in Figure 6).

The user can select two restricted situations, which correspond to the Case 2 and Case 3 mentioned above. If the user push the "Limited Shapes" button, the system responds the barometers on the basis of the situation of Case 2. Therefore, the user can confirm the effectiveness of eliminating some shapes which are easily misrecognized shapes. Similarly, the effectiveness of restriction of rotation ranges is given by pushing the "Limited Ranges" button.

5 CONCLUSIONS

This paper discussed about hand shape selection. When a new hand-gesture based HCI system is designed, a designer can easily know what the system-friendly hand shape is. To develop a supporting system for hand shape selection, 37 candidates of hand shape were captured with various rotations. Then, all the hand shapes were evaluated by several methods including how easily the shape is distinguished from others, how correctly the shape is recognized and so on. Besides, such evaluations were performed under two situations; limited shapes and limited rotations.

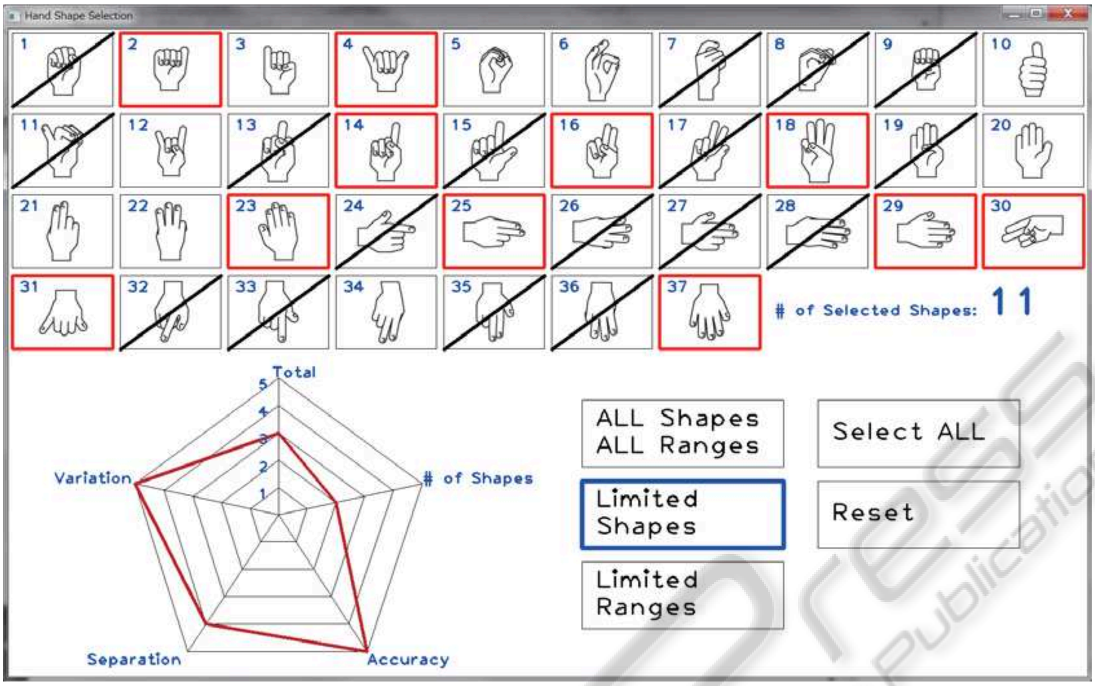

Figure 6: User interface of supporting system for hand shape selection.

Through the system, the user can try to find common ground between "user-friendly" and "system-friendly" hand shapes. We are now investigating whether the selected shapes outperform randomly selected ones by using several recognition strategies, e.g. LSH based approach, and Randomized Trees based approaches and so on. Future work includes developing another interactive system for motion issue of hand gesture recognition.

REFERENCES

Chen, Q., Georganas, N. D., and Petriu, E. M. (2008). Hand gesture recognition using haar-like features and a stochastic context-free grammar. *Ieee Transactions On Instrumentation And Measurement*, 57(8):1562–1571.

de La Gorce, M., Paragios, N., and Fleet, D. J. (2008). Model-based hand tracking with texture, shading and self-occlusions. In *CVPR08*, pages –1–1.

Martin, J. and Crowley, J. L. (1997). An appearance-based approach to gesture-recognition. In *Proceedings, ICIAP*, pages 340–347.

Wachs, J. P., Kolsch, M., Stern, H., and Edan, Y. (2011). Vision-based hand-gesture applications. *Commun.* ACM, 54:60–71.

Wang, C.-C. and Wang, C.-C. (2007). Hand posture recognition using adaboost with sift for human robot interaction. In *Proceedings of the International Conference on Advanced Robotics (ICAR07)*, Jeju, Korea.

Zabulis, X., Baltzakis, H., and Argyros, A. (2009). *Vision-based Hand Gesture Recognition for Human-Computer Interaction*, pages 1–30. Lawrence Erlbaum Associates, Inc. (LEA).