Hand Shape Classification with a Wrist Contour Sensor (Comparison of Feature Types and Observation of Resemblance among Subjects)

Rui Fukui, Masahiko Watanabe, Masamichi Shimosaka and Tomomasa Sato

Abstract Hand gesture can express rich information. However, existing hand shape recognition methods have several problems. In order to utilize hand gesture in a home automation, we have focused on "wrist contour", and have developed a wrist-watch-type device that measures wrist contour using photo reflector arrays. In this paper, we try on two challenges: the first is improvement of the hand shape recognition performance, and the second is making clear the effect of personal difference and finding a key to overcome the difference. We collect wrist contour data from 28 subjects and conduct two kinds of experiments. As for the first challenge, three different feature types are compared. The experimental results extract several important contour statistics and the classification rate itself is also improved by introducing multiple subjects' data for training. As for the second challenge, we compose a resemblance matrix to evaluate resemblance among subjects. The results indicate that training data selection is important to improve the classification performance, especially when we don't have time to collect enough training data for a new user.

1 Introduction

Gesture recognition is getting more popular for home use. Popular gesture devices such as Microsoft Kinect(R) use body movement (i.e. arm movements) as an input. They are specialized in recognition of large body movements, that means they are not good at monitoring small body movements (e.g. hand motion) because of the resolution and accuracy limitations. Even though hand shape can express rich information with small movements, lack of concise hand shape recognition method prevents us from utilizing the beneficial information[8].

The University of Tokyo, 7-3-1 Hongo Bunkyo-ku Tokyo JAPAN, e-mail: {fukui, masahiko, simosaka}@ics.t.u-tokyo.ac.jp, tomomasasato@jcom.home.ne.jp



Rui Fukui, Masahiko Watanabe, Masamichi Shimosaka, and Tomomasa Sato

Therefore, we propose a novel method that realizes hand shape recognition with only a wrist-watch-type device. Such recognition method enables us to realize many applications: remote control of home electronics, gaming interface and so on (Fig. 1).



Fig. 1 Application images: Hand gesture control of displays and recognition of ball grip in baseball game.

1.1 Hand shape recognition methods

There are three major hand shape recognition methods.

- 1. Data glove [1][2]: A user attaches a glove-type device with bend sensors at each finger joint. This method realizes high recognition performance, but the glove restricts hand movements and attaching a glove is a little troublesome for daily activities.
- 2. Camera vision [3][7][6]: This system recognizes hand shapes using image processing. If a camera acquires an image from appropriate direction, it exerts high performance, but it highly depends on the relative position and direction of user's hand and the camera.
- 3. Electromyogram (EMG) [9][11]: With signals from electrode attached to user's arm, it detects myoelectric potential and estimates hand movements. EMG sensor can realize a wearable system. Hence, it does not restrict user position and direction. However, it needs many initial configuration and calibration processes to realize sufficient performance: a user has to clean up his or her skin, the sensor is very naive to attached point, and a recognition system needs a lot of calibration data.

When introducing a hand gesture recognition device into a home, these methods have three problems: influence on activity, complex initial configuration, and insufficient recognition performance.

To overcome these problems, we focus on "wrist contour" We designate a wrist cross-section contour as a wrist contour. The wrist contour has various shapes beHand Shape Classification with a Wrist Contour Sensor

cause finger movements are induced by activities of tendons and muscles near the wrist as shown in Fig. 2.

As a similar approach, Rekimoto[10] developed a wrist-watch-type device with capacitive sensors measuring wrist surface in three points, and recognized two hand shapes (grasping and pointing). In our previous work[5], we developed a wrist-watch-type device (Fig. 3) with two array of 75 photo reflectors, and conducted hand shape classification experiments. The performance of eight hand shapes' classification was 73.2% (training data include subject's own data), or 47.8% (training data exclude subject's own data). The value is not sufficient for practical use, especially personal difference is the bottom neck of the performance.



Fig. 2 Principle of wrist contour variation with hand shape. Fig. 3 The wrist-watch-type device.

1.2 Challenges

In this paper, we try on two challenges: the first is improvement of classification performance, and the second is making clear the effect of personal difference and finding a key to overcome the difference.

As for the first challenge, the wrist contour data differ slightly in attached conditions even in the same subject. Besides, even if the measuring device is almost appressed to the wrist, slippage (moving in a radial or circumferential direction) may occur. Therefore, we need to investigate a robust normalizing procedure that can extract stable characteristics and need to design more solid feature extraction process.

Regarding the second challenge, we want to utilize other subjects' data that is acquired previously, and this approach can release a new user from collecting numerous training data.

2 Approaches for Efficient Utilization of Wrist Contour Data

In order to measure wrist contour precisely, we developed a wrist contour measuring device. From the data collected with the device, features are extracted for hand shape classification methods. The key elements for our approach are described in this section.

2.1 Wrist Contour Measuring Device

The device consists of two parts: the wrist-watch-type measurement part, and the battery and control part (Fig. 4). The measurement part has a measurement band; a flexible band with photo reflectors (infrared-light distance sensors). The band can measure distances between the band and surface of the wrist using photo reflector arrays. The band has two arrays and each array has 75 photo reflectors. With a wireless module mounted in battery and control part, the device communicates with PC wirelessly.

The basic specifications are as follows. Please refer to our previous work [5] for the details of the device.

- Measurement area: ~185mm
- Measurement pitch: 2.5mm
- Distance resolution: 0.1mm (~3.5mm range)
- Sampling rate: 10Hz

2.2 Feature Extraction Process

Wrist contour raw data examples are shown in the upper side of Fig. 5. There are small difference among the raw data of hand classes, therefore feature extraction process is essential. We prepare two potential feature types.

One feature type is "normalized contour data". Because each muscle and tendon is different in thickness, each sensor element has different variation range of distance. The process samples the maximum and minimum distance for each sensor element, and normalizes distance data into 0 to 1 (Fig. 5). With this process, small variations can be emphasized. On the other hand, the slippage of the band might be great noise.

Another feature type is "contour statistics". They are statistics from wrist contour distance data, such as sum of distances, maximum distance, histograms and so on (Fig. 6). Each statistics are normalized by calibration data (wrist contour data of Fist and Open hand). With this approach, we try to overcome slippage or personal differences.



Fig. 4 Composition of wrist contour measuring device. The measurement part is connected to the battery and control part by a wire. The Fixing band assists the attachment of the measurement band and reduces slippage.



Fig. 5 Example of wrist contour data normalization.

Fig. 6 Example of statistics from wrist contour data.

2.3 Classification Method

As classification methods, k-Nearest Neighbor (k-NN) method and AdaBoost method are used.

In k-NN method, the test data is labeled by votes of k nearest samples. All Euclidean distances between the test sample and training samples in the feature space are calculated, and nearest k training samples have right to vote.

AdaBoost is a kind of boosting method, which makes some weak learners (in our implementation, decision stumps) [4]. The test data is labeled by weak learners' weighted votes. Weights on weak learners are tuned to fit to training data in the training process.

2.4 Personal Resemblance

Resemblance among subjects is the main topic to be investigated. That is because when other subjects' data are used as training data, the recognition performance differs drastically according to the combination of test subject and training subjects. We try to observe the resemblance among subjects by examining the relationships between classification performance and combination of subjects.

3 Experiments

We collected wrist contour data from 28 subjects. With the data, two kinds of hand shape classification experiments are conducted: (1) Comparison of two feature types using three training data groups, and (2) Evaluation of resemblance among subjects.

3.1 Wrist Contour Data Collection

We collected wrist contour data from 28 subjects, male and female of 20's to 50's. The arm posture and wrist pronation is fixed as shown in Fig. 7 because a wrist contour varies with wrist pronation. Data collecting procedure is as follows:

Step 1 The measurement part is attached on the wrist in rough alignment.

Step 2 A display shows a hand shape illustration to the subject, and the subject imitates the hand shape, and then wrist contour data is recorded.

Step 3 After recording wrist contour data of all six hand shapes (one set, shown in Fig. 8), the measuring device is taken off.

Step 4 Repeat Step $1 \sim 3$ six times for each subject.

Finally, 1008 wrist contour data were collected; 28 subjects \times 6 wrist contour data sets \times 6 hand shapes. Examples of wrist contour raw data are shown in Fig. 9.



Fig. 7 The posture **Fig. 8** Hand shape images when collecting wrist of six hand classes data. (one set).



3.2 Classification Experiment (1) Comparison of Two Feature Types

In order to examine robustness of the described feature extraction processes, two types and one derived type of features are compared.

3.2.1 Setup of Classification Experiment (1)

Feature types to be compared are as follows.

- A. Normalized contour data 45 dimensions: each contour data are linearly converted to the smallest wrist contour size.
- B. Contour statistics 92 dimensions
- C. Selected contour statistics 5 dimensions: This feature type is derived from feature type B. When classifying from contour statistics using k-NN method, useful features might be hidden by other less useful features. Also, when classifying using AdaBoost, large number of features might cause over-fitting. Therefore we conduct another experiment of using only five contour statistics that have large separation metrics. The separation metrics is calculated as between-class variance divided by in-class variance.

Two classification methods are used: k-NN method and AdaBoost method. When using AdaBoost, the system produces six one-versus-the-rest classifiers and an output class is determined by maximum output value in six classifiers.

Therefore, the number of combinations of feature types and classification methods are $3 \times 2 = 6$.

As for training data, three groups are prepared.

Group 1	Subject's own data 5 sets.
Group 2	Subject's own data 5 sets and other subjects' 27×6 sets.
Group 3	Other subjects' 27×6 sets.

The classification performance is evaluated by classification rate (number of correctly classified samples / number of inputted samples $\times 100$ [%]).

3.2.2 Result of Classification Experiment (1)

Result of classification experiment (1) is shown in Fig. 10. The table explains the result in each training data groups. As expected, the combination of feature type B and k-NN method exerts low score in all groups.

In group 1 (the subject's own data), feature type A exerts higher performance in classification rate than feature type B and C. Incidentally, results of experiments using AdaBoost (especially feature type A and B) exert lower performance. This is because there are not enough training data and that causes over-fitting. In group 2 (the subject's own data and other subjects' data), no remarkable difference among three feature types is observed. In group 3 (other subjects' data), feature type B and C exert better performance than feature type A.

Tbl. 1 shows the confusion matrix of one experiment (B. Statistics + AdaBoost). Most hand shapes are correctly classified on some level, however, confusions among some similar classes (e.g. 1F and LF) are observable.



Fig. 10 Result of classification experiment (1).

The insights from the results are as follows. First, useful features are different according to training data group. For instance, when using subject's own data as training data, the combination of feature type A (normalized contour data) and k-NN classification method exerted relatively high performance of 90.1% classification rate. On the other hand, when using other subjects' data as training data, the performance is improved using feature type B (contour statistics); classification rate is 77.9% using AdaBoost. Second, the weights of AdaBoost weak learners indicates

8

that sum of distances, maximum monotone increasing value and sum of differences of histogram are useful contour statistics.

As a result, it can be said the performance is improved: with subject's own data, 90.1% in classification of six hand shapes is thought to be enough for use. However, the classification rate of 77.9% with other subjects' data is not enough for practical use, so we should consider how to use other subjects' data more efficiently.

3.3 Classification Experiment (2) Evaluation of Resemblance among Subjects

In order to examine resemblance in wrist contours among subjects, classification experiment is conducted in one-on-one combinations of training and test data.

3.3.1 Setup of Classification Experiment (2)

Classification rates are used for evaluation of resemblance among subjects. We configure the resemblance evaluation matrix (28 subjects \times 28 subjects). To make the matrix, all combination (28 \times 27) classification experiments are conducted. For each experiment, training data only include one subject data. An element (S1, S2) of the resemblance evaluation matrix is classification rate when using S1's data as test data, and S2's data as training data.

Two feature types (A and C) are used and k-NN method is used as classification method. This is because when classifying with AdaBoost, learners are over-fitted to small number of training data.

3.3.2 Result of Classification Experiment (2)

One result of classification experiment (2) is shown in Tbl. 2; the result of the experiment using feature type C (selected contour statistics). The row represents the test subject and the column represents the training subject. For example, when the test subject is subject 3 and the training subject is subject 9, the classification rate is 80.6% (the green box).

The average classification rate of all experiments is 59.8%. The rate of more than 80% elements is 11.2% (85/756), and the rate of less than 40% elements is 12.8% (97/756).

Meanwhile, the result of experiment using feature A is as follows: the average classification rate of all experiments is 54.3%, the rate of more than 80% elements is 5.6% (42/756), and the rate of less than 40% elements is 20.2% (153/756). In addition, each element of the resemblance evaluation matrixes is far different from the one of the result using feature C.



Several insights can be drawn from the results. First, the fact that the rate of useful combinations (more than 80%) is 11.2% and the rate of useless combinations (less than 40%) is 12.8% indicates the necessity of a training data selection process. Second, the resemblance evaluation matrix of feature type C and the one of feature type A are far different, that means the definition of resemblance deeply depends on the type of features. Therefore the training data selection processes need to be designed respectively to the feature type. Third, when examining each subject, there are some subjects that exert low classification rate with any combinations, so we need to collect more data to fill the "lack" of current data.

When applying results of experiment (2) to improve classification performance, some approaches can be candidate. One approach is making groups of highly resemble subjects, and then training data or classifiers are selected according to the group that new subject belongs to. It is important to determine what model to use as group, and number of groups, and design of group classifiers. Another candidate approach is like a kind of filtering that selects each subject to use as training data [12].

4 Conclusion and Future Work

We focus on wrist contour for hand shape recognition in order to overcome problems of existing recognition methods. As a measurement device, we developed a wrist-watch-type device with photo reflector arrays. Using the device, wrist contour data of six hand shapes are collected from 28 subjects. With the collected data, Hand Shape Classification with a Wrist Contour Sensor

we conducted several experiments for comparison of feature types, and evaluated resemblance among subjects.

Through classification experiment (1), it can be said the performance is improved. However, the classification rate of 77.9% with other subjects' data is not enough for practical use and more efficient usage of other subjects' data is essential. Through classification experiment (2), the need of training data selection process are confirmed. We should try on some method of grouping or selecting training data.

Our future works are as follows. In terms of design, it is necessary to downsize the device and integrate all functions to a wrist-watch-type device. Additionally, we need to tackle the problem of wrist pronation changes by redesigning hardware and software. In terms of recognition performance, we need to try some approaches such as finding other new statistics or using another classification methods to improve performance. As mentioned above, an efficient training data selection process should be designed to utilize other subjects' data as training data.

References

- 1. CyberGloveSystems.
 - http://www.cyberglovesystems.com/.
- Deller, M., Ebert, A., Bender, M., Hagen, H.: Flexible gesture recognition for immersive virtual environments. In: Proceedings of Tenth International Conference on Information Visualization., pp. 563–568, (2006)
- Erol, A., Bebis, G., Nicolescu, M., Boyle, R.D., Twombly, X.: A review on vision-based full DOF hand motion estimation. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, p. 75, (2005)
- Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. In: Proceedings of the Second European Conference on Computational Learning Theory, pp. 23–37, (1995)
- Fukui, R., Watanabe, M., Gyota, T., Shimosaka, M., Sato, T.: Hand shape classification with a wrist contour sensor: Development of a prototype device. In: Proceedings of ACM Conference on Ubiquitous Computing, pp. 311–314, (2011)
- Garg, P., Aggarwal, N., Sofat, S.: Vision based hand gesture recognition. World Academy of Science Engineering and Technology, pp. 972–977, (2009)
- 7. Hoshino, K., Tamaki, E., Tanimoto, T.: Copycat hand robot hand imitating human motions at high speed and with high accuracy. Advanced Robotics, Vol. 21, pp. 1743–1761, (2007)
- Huang, T.S., Pavlovic, V.I.: Hand gesture modeling, analysis, and synthesis. In: Proceedings of IEEE International Workshop on Automatic Face and Gesture Recognition, pp. 73–79, (1995)
- Nagata, K., Adno, K., Yamada, M., Magatani, K.: A classification method of hand movements using multi channel electrode. In: Proceedings of IEEE-EMBS. Annual International Conference of the Engineering in Medicine and Biology Society, pp. 2375–2378, (2005)
- Rekimoto, J.: GestureWrist and GesturePad: Unobtrusive wearable interaction devices. In: Proceedings of the 5th IEEE International Symposium on Wearable Computers, pp. 21-27, (2001)
- Yoshikawa, M., Mikawa, M., Tanaka, K.: Real-time hand motion estimation using EMG signals with support vector machines. In: Proceedings of SICE-ICASE. International Joint Conference, pp. 593–598, (2006)
- Yu, K., Tresp, V., Yu, S.: A nonparametric hierarchical bayesian framework for information filtering. In: Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 353–360, (2004)