We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



18

Copycat Hand - Robot Hand Generating Imitative Behaviour at High Speed and with High Accuracy

Kiyoshi Hoshino
University of Tsukuba
Japan

1. Introduction

In recent years, robots demonstrating an aping or imitating function have been proposed. Such functions can estimate the actions of a human being by using a non-contact method and reproduce the same actions. However, very few systems can imitate the behaviour of the hand or the fingers (Bernardin et al., 2005). On the other hand, reports in Neuroscience state that mirror neurons (Rizzolatti et al., 1996; Gallese et al., 1996), which participate in the actions imitated by chimpanzees, are activated only for actions of the hand and fingers, such as cracking peanut shells, placing a sheet of paper over an object, tearing the sheet into pieces, etc.. Moreover, since human beings perform intelligent actions using their hands, it is important to attempt to artificially imitate the actions of hands and fingers in order to understand the dexterity and intelligence of human hands from an engineering viewpoint. The object actions in the case where one "looks at an action performed by others and imitates it" include not only grasping or manipulating objects but also the actions involving "imitating shapes and postures of hands and fingers" of others such as sign language and dancing motions. The latter two types of actions are often more complicated and require dexterity as compared to the former actions.

In the action of imitating "the shape of hands and fingers" such as sign language, it is essential to estimate the shape of the hands and fingers. Furthermore, as compared to the imitation of the actions of the lower limbs, estimating the shape of the hands is significantly more important and difficult. In particular, human hands, which have a multi-joint structure, change their shapes in a complicated manner and often perform actions with self-occlusion in which the portion of one's own body renders other portions invisible. In the case of an artificial system, we can utilize a multi-camera system that records a human hand for imitative behaviours by surrounding it with a large number of cameras. However, all the animals that mimic the motions of others have only two eyes. To realize the reproduction of the actions of hands and fingers by imitating their behaviour, it is desirable to adopt a single-eye or double-eye system construction.

To roughly classify conventional hand posture estimation systems, the following two types of approaches can be used. The first approach is a 3D-model-based approach (Rehg & Kanade, 1994; Kameda & Minoh, 1996; Lu et al., 2003) that consists of extracting the local characteristics, or silhouette, in an image recorded using a camera and fitting a 3D hand model, which has been constructed in advance in a computer, to it. The second approach is a 2D-appearance-based approach (Athitos & Scarloff, 2002; Hoshino & Tanimoto, 2005) that

Open Access Database www.i-techonline.com

Source: Humanoid Robots, New Developments, Book edited by: Armando Carlos de Pina Filho
ISBN 978-3-902613-02-8, pp.582, I-Tech, Vienna, Austria, June 2007

consists of directly comparing the input image with the appearance of the image stored in a database. The former is capable of high-accuracy estimations of the shape, but it is weak against self-occlusion and also requires a long processing time. The latter can reduce the computation time; however, if 3D changes in the appearance of hands are not an issue, which also include the motions of the wrist and the forearm, a large-scale reference database is required, and it becomes difficult to control the robot hand by means of imitation. However, if the fundamental difficulty in the estimation of the hand posture lies in the complexity of the hand shape and self-occlusion, the high-accuracy estimation of the shape will become theoretically possible; this requires the preparation of an extensive database that includes hand images of all the possible appearances with complexity and self-occlusion. The feasibility of this approach depends on the search algorithm used for rapidly finding similar images from an extensive database.

Therefore, in this study, we aim at realizing a system for estimating the human hand shape capable of reproducing actions that are the same as those of human hands and fingers to a level reproducible using a robot hand at high speeds and with high accuracy. With regard to the processing speed, we aimed to achieve short-time processing of a level enabling the proportional derivative (PD) control of a robot. With regard to the estimation accuracy, we aimed at achieving a level of estimation error considered almost equivalent to that by the visual inspection of a human being, namely, limiting the estimation error in the joint angle to within several degrees. In addition to those two factors, dispersion in the estimation time due to the actions of human hands and fingers causes difficulties in the control of a robot. Therefore, we also aimed to achieve uniformity in the estimation time. In particular, in order to conduct high-speed searches of similar data, we constructed a large-scale database using simple techniques and divided it into an approximately uniform number of classes and data by adopting the multistage self-organizing map (SOM) process (Kohonen, 1988), including self-multiplication and self-extinction, and realized the high-speed and high-accuracy estimation of the shape of a finger within a uniform processing time. We finally integrated the hand posture estimation system (Hoshino & Tanimoto, 2005; 2006) with the humanoid robot hand (Hoshino & Kawabuchi, 2005; 2006) designed by our research group; our final system is referred to as the “copycat hand”.

2. System construction

2.1 Construction of a large-scale database

First, we conducted measurements of the articular angle data (hereafter referred to as the “key angle data”). In the present study, we determined the articular angle at 22 points per hand in the form of a Euler angle by using a data glove (Cyberglove, Virtual Technologies). For the articular data in the shape of a hand that requires a high detecting accuracy, especially in the search for similar images, a large number of measurements were made. Furthermore, we generated CG images of the hand on the basis of the key articular angle data. The CG editing software Poser 5 (Curious Labs Incorporated) was used for the generation of images.

Second, from the two key angle data, we interpolated a plurality of the articular angle data in optional proportions. The interpolation of the articular data is linear. Moreover, we also generated the corresponding CG images of the hand on the basis of the interpolated data. This operation enables an experimenter equipped with the data glove to obtain CG images of the hand in various shapes with the desired fineness without having to measure all the

shapes of the hand. Fig.1 shows a schematic chart of the interpolation of the articular angle data and the CG images of the hand. Furthermore, Fig.2 shows an example of the interpolated CG images of the hand. This figure represents an example of a case where the articular angle was measured at three different points in time for the actions of changing from 'rock' to 'scissors' in the rock-paper-scissors game, and the direct generation of CG and the generation of CG using interpolation were made from two adjoining data. In both these figures, the three images surrounded by a square represent the former, while the other images represent the latter.

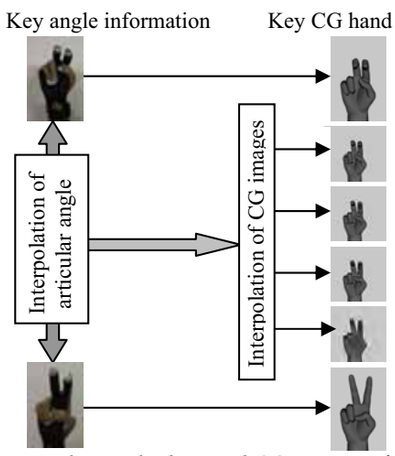


Fig. 1. Interpolation of the articular angle data and CG images of the hand.

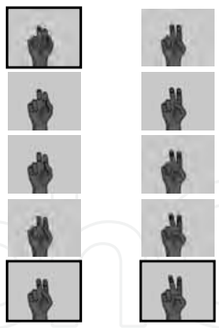


Fig. 2. Examples of the interpolated CG images of the hand.

Third, we added the data describing the differences among individuals. Because of the differences that exist among individuals (as shown in Fig.3), a wide variety of data is required for a database intended for searching similar images. For example, in the hand shape representing 'rock' in the rock-paper-scissors game, a significant difference among individuals is likely to appear in (1) the curvature of the coxa position of the four fingers other than the thumb and (2) the manner of protrusion of the thumb coxa.

Moreover, differences are likely to appear in (3) the manner of opening of the index and the middle finger and (4) the standing angle of the reference finger in the 'scissors' shape, and also in (5) the manner of opening and (6) the manner of warping, etc. of the thumb in the 'paper' shape. In order to express such differences among individuals in the form of the CG hand, we need to adjust the parameters of the length of the finger bone and the movable articular angle; therefore, we generated the CG images of hands having differences among individuals on the basis of the articular angle data obtained by the procedure described above. Fig.4 indicates an example of the additional generation of the CG hand in different shapes. In the figure, the X axis shows CG hands arranged in the order starting from those with larger projections of the thumb coxa, while the Y axis represents those with larger curvature formed by the coxa of the four fingers other than the thumb, respectively.

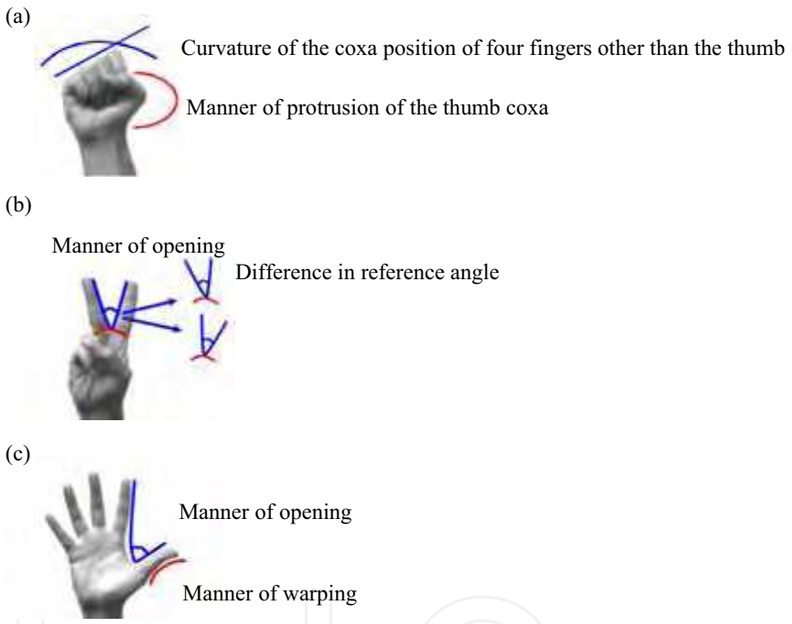


Fig. 3. Examples of the differences among individuals.

By performing the first to third steps mentioned above, we generated a total of 15,000 CG hand images using this system.

Then, the resolution was changed. Although the CG image generated this time had a resolution of 320 x 240 pixels, a substantial calculation time is required in order to estimate the posture and for applying various image processing techniques. In the present study, a reduced resolution of 64 x 64 was used. The pixel value after the resolution was changed is given by the following expression:

$$gr(i,j) = \frac{1}{r}\sum_{k}\sum_{l} go(i*320/64+k, j*320/64+l) \quad (1)$$

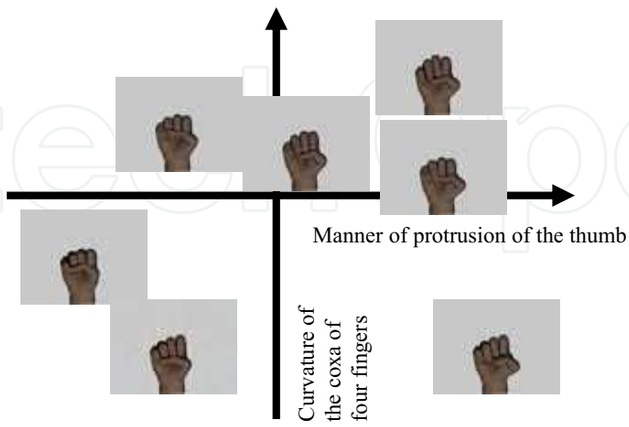


Fig. 4. Examples of the supplemented data of the differences among individuals.

Here, $gr(i,j)$ and $go(i,j)$ are the pixel values at row i and column j after and before altering the resolution, respectively. Here, the calculation has also been vertically conducted with 320 pixels in order to match the aspect ratio since the pixel resolution was altered to 64 x 64. Furthermore, k and l correspond to the row and column, respectively, within the respective regions before changing the resolution, and $r = k \times l$.

Finally, the contour was extracted. Differences exist in the environmental light, colour of human skin, etc. in the input images. The abovementioned factors were eliminated by extracting the contour in order to fix the width and the edge values, and the estimation errors were reduced by reducing the difference between the hand images in the database and in the input data.

2.2 Characterization

In the present study, we used the higher-order local autocorrelational function (Otsu & Kurita, 1998). The characteristics defined using the following expression were calculated with respect to the reference point and its vicinity:

$$x^N(a_1, a_2, \cdots, a_N) = \int f(r) f(r + a_1) \cdots f(r + a_N) dr \tag{2}$$

Here, x^N is the correlational function in the vicinity of the point r in dimension N. Since the pixels around the object point are important when a recorded image is generally used as the processing object, the factor N was limited up to the second order in the present study. When excluding the equivalent terms due to parallel translation, x^N is possibly expressed using 25 types of characteristic quantities, as shown in Fig.5. However, patterns M1 through M5 should be normalized since they have a smaller scale than the characteristic quantities of patterns M6 and thereafter. By further multiplying the pixel values of the reference point for patterns M2 through M5 and by multiplying the square of the pixel value of the reference point for pattern M1, a good agreement with the other characteristic quantities was obtained. In the present study, an image was divided into 64 sections in total – 8 x 8 each in the vertical and lateral directions - and the respective divided images were represented by 25 types of characteristic quantities using the higher-order local autocorrelational function.

Therefore, a single image is described using the characteristic quantities of 25 patterns x 64 divided sections. The image characteristics of the CG hand and the joint angle data were paired as a set for preparing the database.

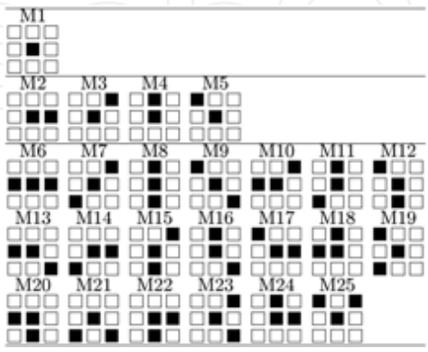


Fig. 5. Patterns of the higher-order local autocorrelational function.

2.3 Self-organization of the database

If the database prepared in the preceding sections is directly used for searching, it increases the search time together with a larger database. Hence, we intend to narrow the search space by clustering data with similar characteristics in the database. For example, sorting by using a dichotomizing search may be feasible for ordinary data; however, in the case where the characteristics range over multiple dimensions, a limitation is that the number of searches during a retrieval becomes the same as that in the total search. Therefore, we constructed a database using Kohonen's SOM (Kohonen, 1988).

Each database entry has a joint angle and a number of image characteristics; however, only the image characteristics are used in the search during estimation. There is a possibility that there exist data that have similar characteristics but significantly different joint angles; such data may be included in the same class if the classification is made on the basis of the characteristics during the self-organization of the database. On the other hand, there also exist data having significantly different characteristics, although the joint angles are similar. Therefore, we performed self-organization for both these types of data and conducted preliminary experiments; the obtained results are listed in Table 1. The mean value of the errors and the standard deviation are the values for the middle finger. The data for the other fingers are omitted from the table since they exhibited similar tendencies. Degree is used as the unit of the mean value of the errors and the standard deviation. As shown in the table, the case of self-organization on the basis of characteristics yielded better results. Consequently, we performed data clustering using self-organization on the basis of characteristics in the present study.

	processing time [ms]	mean error [degree]	standard deviation
joint angle	0.842	0.792	6.576
characteristics	0.764	0.373	5.565

Table 1. Performance of self-organization on the basis of joint angles and characteristics in the preliminary experiment.

First, we prepared classes having the representative angle, representative number of characteristics and neighbourhood class information as classes in the initial period. For the initial angles and the number of characteristics, random numbers in the range of 0 to 1 were used. With regard to the neighbourhood class information, we calculated the distance between classes in the angles by using the Euclidean distance and determined classes close to one another in this distance as neighbouring classes; this information was retained as the class number. Although the number of neighbouring classes depends on the scale of the database and the processing performance of the PC, we studied it heuristically in this experiment, and determined classes up to that close to the eighth as the neighbour classes. Next, we calculated the distance in the characteristics between the data and the classes and selected the closest class by using the data in a secondary database. This class will hereafter be referred to as the closest neighbour class. Moreover, the used date will be considered as those belonging to the closest neighbour class. The representative angle and representative number of characteristics of the closest neighbour class were renewed by using the expression below so that they may be placed closer to the data.

$$\begin{aligned} CA_{ij} &= CA_{ij} - \alpha(CA_{ij} - DA_{rj}) \\ CF_{ij} &= CF_{ij} - \alpha(CF_{ij} - DF_{rj}) \end{aligned} \tag{3}$$

where CA_{ij} denotes the representative angle j of class i; DA_{rj}, the angle j of data r; CF_{ij}, the representative number of characteristics j of class i; DF_{rj}, the representative number of characteristics j of data r; and a, the coefficient of learning.

In this experiment, a was heuristically determined as 0.0001. Next, a similar renewal was also made in the classes included in the neighbour class information of the closest neighbour class. However, their coefficient of learning was set to a value lower than that of the closest neighbour class. In the present study, it was heuristically selected as 0.01. This was applied to all the data in the primary database. In order to perform self-organization, the abovementioned operation was repeated until there was almost no change in the representative angle and the representative number of characteristics of the class.

Narrowing and acceleration of the search process can be realized to some extent, even if the database is used without self-organization. However, if such a database is used, dispersion is observed in the amount of data included in each class, thereby inducing dispersion in the processing time. Therefore, we intended to avoid the lack of uniformity in the processing time by introducing an algorithm for self-multiplication and self-extinction during self-organization. After selecting the class of adherence for all the data, we duplicate the classes that contain an amount of data exceeding 1.5 times the ideal amount. In addition, we deleted the classes containing an amount of data no more than one-half the ideal amount of data. Therefore, the amount of data belonging to each class was maintained within a certain range without significant dispersion, and the processing time was maintained within a certain limit, irrespective of the data in the class that was used for searching during the estimation. In case the algorithm for self-multiplication and self-extinction is introduced, a change is produced in the relationships among the classes, which remains unchanged in ordinary self-organization, making it necessary to redefine the relationships among the classes. Therefore, we newly prepared the neighbour class information by a method similar to that used during initialization in which we duplicated and deleted the classes.

Estimations made by using a database obtained in this manner can considerably increase the search speed as compared to the complete search of ordinary data. However, considering further increases in the database and acceleration of the searches, the database clustering was

performed not only in a single layer but also in multiple layers. Fig. 6 shows the schematic structure of the multiple layers. The class obtained with the aforementioned processing is defined as the second-layer class and is considered as data. A third-layer class is prepared by clustering the second-layer classes as the data. The third-layer class is prepared by following the same procedure as that used in the preparation of the second-layer class. Further, a fourth-layer class is prepared by clustering the third-layer classes. The lesser the amount of data in one class (or the number of classes in the lower layers), the higher the layer in which clustering can be performed. However, to absorb the dispersion of data, etc., it is preferable to prepare classes having an amount of data with a certain volume. Table 2 lists the results of the preliminary experiment in which clustering was performed by setting the amount of data in a class at 5, 10 and 20. Although the search time is reduced if the clustering is performed with a small amount of data, the estimation accuracy also reduces accordingly; therefore, we set an ideal amount of data as 10 in the present study as a trade-off between the two parameters.

The clustered database obtained using the abovementioned operation was termed as a tertiary database. This tertiary database will hereafter be simply referred to as the database. In this system, we finally constructed a database comprising 5, 10 and 10 classes in order from the upper layers, where each class has approximately 10 data items.

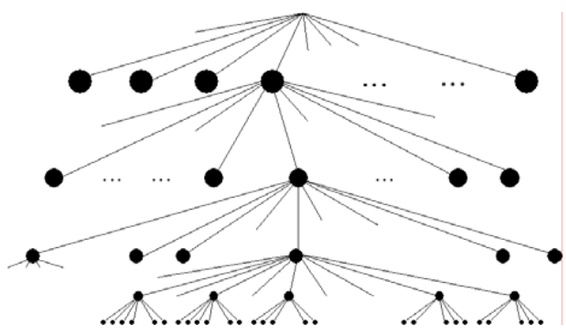


Fig. 6. Schematic structure of a database with multiple layers.

	processing time [ms]	mean error [degree]	standard deviation
5	0.656	-0.035	5.868
10	0.764	0.373	5.565
20	1.086	0.145	5.400

Table 2. Performance according to the amount of data in a class in the preliminary experiment.

2.4 Search of similar images

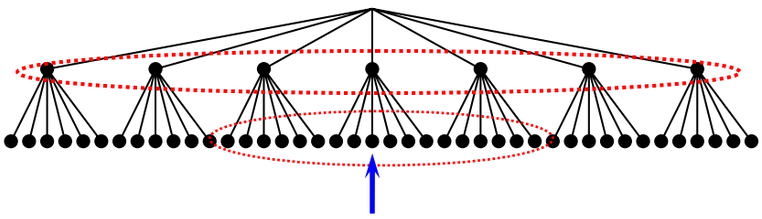
During estimation, sequential images were acquired using a high-speed camera. In a manner similar to the preparation of the database, image processing techniques were applied to these images to obtain their characteristic quantities. By comparing each quantity with that in the database by means of a processing technique described later, the joint angle information that formed a pair with the most similar image were defined as each result was estimated.

To estimate the similarity at the first search, the distance was calculated by using the characteristic quantity for all classes in the database. The calculation was performed by simply using the Euclidean distance that is derived using the expression below:

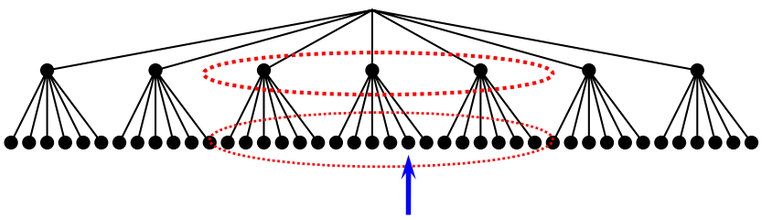
$$E_r = \sum_{i=1}^{25*n} (x_{ri} - x_{ti})^2 \quad (4)$$

Here, both x_{ri} and x_{ti} are characteristic quantities i with the higher-order local autocorrelational functions of the class r and at the time t, respectively. The class that minimizes E_r was selected as the most vicinal class at time t. With respect to the affiliated data of the most vicinal class and all the vicinal classes of the most vicinal class, the distances from the characteristic quantities obtained from the image were calculated using expression (4). At each instance, the angle of the data with the shortest distance was regarded as the estimated angle. From the second search, the distance was not calculated by using the characteristic quantity for all the classes in the database. Instead, only the vicinal classes of the most vicinal class and the affiliated data were selected as the candidates for the search according to the histories at t-1, as shown in Fig.7.

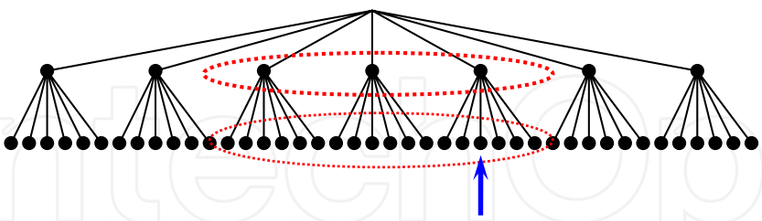
(a) at first search: all classes are candidates for the search.



(b) from second search, the vicinal classes of the most vicinal class are candidates.



(c) if the result moves to and affiliates with another class,



(d) then, the search space and candidate classes moves.

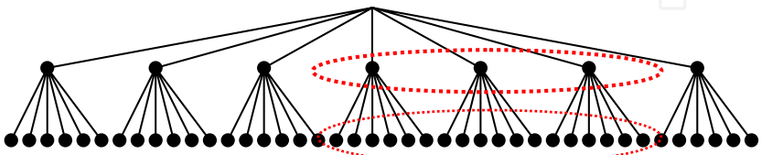


Fig. 7. Differences in the search spaces between the first search and the succeeding searches.

3. Experiment of posture estimation

3.1 Methods and procedures

In order to verify the effectiveness of this system, the actual images were subjected to experimental estimation. A subject held up a hand at a position approximately 1 m in front of the high-speed camera and moved the fingers freely provided the palm faced the camera. A slight motion of the hand was allowed in all the directions provided the hand was within the field angle of the camera. We employed a PC (CPU: Pentium 4, 2.8 GHz; main memory: 512 MB) and a monochromatic high-speed camera (ES-310/T manufactured by MEGAPLUS Inc.) in the experiments.

3.2 Results and discussions

Fig.8 shows the examples of the estimation. Each estimated result plotted using the wireframe model was superimposed on the actual image of a hand. It is evident that the finger angles have possibly been estimated with a high precision when the hand and fingers were continuously moved. It was verified that the estimation could be performed, provided the hand image did not blend into the background, even if the illuminating environment was changed.

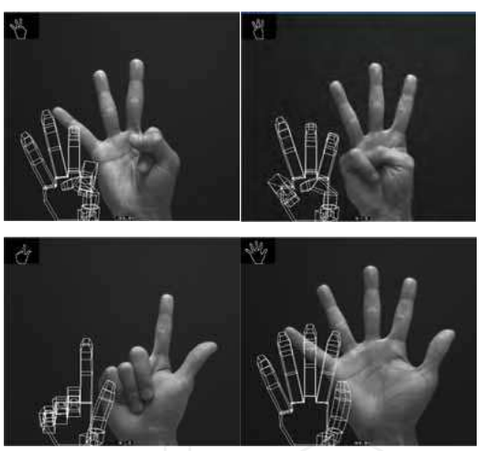


Fig. 8. Captured hand images and the results of the hand posture estimation.

For the purpose of a quantitative assessment of the system, the measured and estimated values have to be compared. However, in an ordinary environment using this system, it is impossible to acquire the measured values of the joint angle information from the human hand and fingers moving in front of the camera. Consequently, we performed the estimation experiment by wearing the data glove and a white glove above it. The results are shown in Fig.9, which reveals the angular data measured using the data glove and the estimated results. Fig.9(a) shows the interphalangeal (IP) joint of the thumb; Fig.9(b), the abduction between the middle and ring fingers; and Fig.9(c), the proximal interphalangeal (PIP) joint of the middle finger. The state where the joint is unfolded was set as 180 degrees. The system at this time operates at more than 150 fps and thus enables realtime estimation.

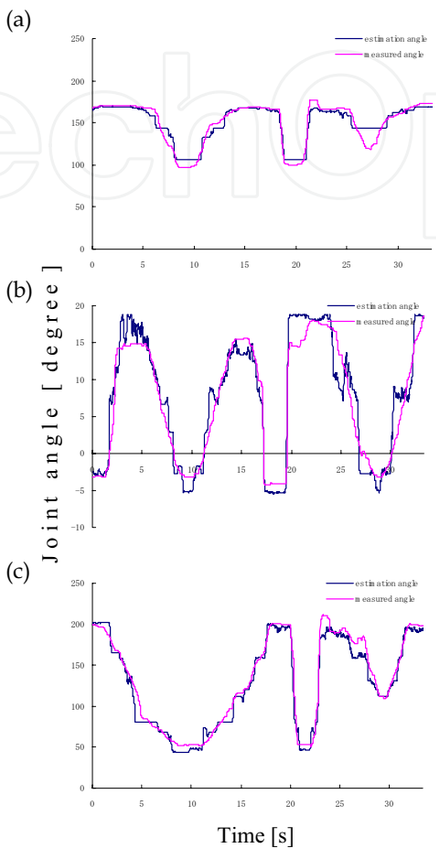


Fig. 9. Examples of the joint angle data measured using the data glove and the estimated results.

As evident from the figure, the standard deviation of the errors in the estimated angles was 4.51 degrees when we avoided the fluorescent light and used the angular data obtained by means of the data glove as the actual values; the results obtained did not have highly precise numerical values. We observed a trend of poor estimations, particularly for parts with little variation in the image (for example, the shape of the rock in the rock-paper-scissors game) against the angular variation. This may be expected, considering that a human is performing the figure estimation. In other words, we can hardly observe any difference visually for an angular difference of 10 degrees when each finger has a difference of 10 degrees. Therefore, the errors in this system, which conducts estimation on the basis of the camera image, may be considered as being within the allowable range. On the contrary, it can be observed from this figure that highly precise estimations are made in the region where visual differences are observed, namely, where the image changes significantly with the angular variations and where it is located in between the flexion and the extension.

Next, the comparative experiments were conducted. The difference between the previous experiment and these comparative experiments is that the hand position agrees with or closely resembles the database image since the object for estimation is set by selecting the CG hand image from the database. Consequently, we can determine the expected improvement in the estimating precision when the processing for positioning the input image is integrated into this system. The standard deviation of the errors when estimating the object was set to 2.86 degrees by selecting the CG image from the database, thus allowing very high-precision estimation. It is expected that the estimation error can be reduced to this extent in the future by integrating the processing for correcting the position into this system. Moreover, the processing time for the search, except for the image processing, is 0.69 ms per image. From the viewpoint of precision and processing speed, the effectiveness of the multi-step search using the self-organized database has been proved.

As mentioned above, the estimation error for unknown input images had a standard deviation of 4.51 degrees. Since this is an image processing system, small variations in the finger joints in the rock state of the rock-paper-scissors game will definitely exhibit a minimal difference in the appearance; these differences will numerically appear as a large error in the estimation. However, this error possibly contains calibration errors arising from the use of the data glove, as well as the errors caused by slight differences in the thickness, colour, or texture of the data glove covered with the white glove. Therefore, the output of the data glove or the actual value of the quantitative assessment requires calibration between the strain gauge output and the finger joint value whenever the glove is worn since the joint angle is calculated from a strain gauge worn on the glove. No such calibration standards exist, particularly for the state in which the finger is extended; therefore, the measured angle can be easily different from the indicated value. Even when the estimation is newly calibrated, it is possible that the state of calibration may be different in each experiment. On the other hand, it is not necessary to apply calibration to the second experiment that selects the CG hand image from the database. It is highly possible that this influences the standard deviation value of 4.51 degrees; therefore, it is possible to consider that the standard deviation of the errors lies between 4.51 and 2.86 degrees even if the system has not been subjected to corrective processing for the hand position.

The scheme of the present study allows you to add new data even without understanding the system. Another advantage is that the addition of new data does not require a long time since it is unnecessary to reorganize the database even when several new data items are added; this is because the database can sequentially self-organize itself by using the algorithm for self-multiplication and self-extinction of database classes. Furthermore, it is possible to search the neighbouring classes having angular similarities since each class possesses information about the vicinal classes in this system. This fact can also be regarded as the best fit for estimating the posture of a physical object that causes successive temporal angular variations, such as estimating the posture of the human hand. We attempted to carry out the hand posture estimation when the hand is rotated, although the number of trials was inadequate. Fig.10 shows an example of the result, which suggests that our system functions when a subject is in front of the camera and is rotating his/her hand. A subject can also swing the forearm, and our system can effectively estimate the shape of the fingers, as shown in Fig.11.

The image information and the joint angle information are paired in the database in our system. Once we output the results of the hand posture estimation to a robot hand, the robot can reproduce the same motions as those of the fingers of a human being and mimic them.

Fig.12 shows a dexterous robot hand (Hoshino & Kawabuchi, 2005) imitating the human hand motions without any sensors attached to it. We refer to this integrated system as the "copycat hand". This system can generate imitative behaviours of the hand because the hand posture estimation system performs calculations at high speeds and with high accuracy.



Fig. 10. An example of hand posture estimation using the rotating motion of the wrist.

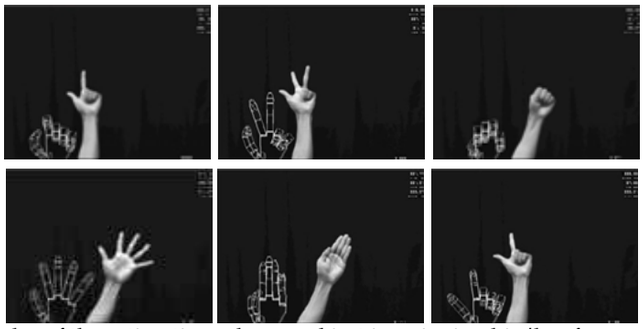


Fig. 11. Examples of the estimation when a subject is swinging his/her forearm.



Fig. 12. The copycat hand can ape and imitate human hand motions at high speeds and with high accuracy.

4. Conclusion

To realize a robot hand capable of instantly imitating human actions, high speed, high accuracy and uniform processing time in the hand posture estimation are essential. Therefore, in the present study, we have developed a method that enables the searching of similar images at high speeds and with high accuracy and the search involves uniform processing time, even in the case where a large-scale database is used. This is achieved by (1) clustering databases having approximately uniform amounts of data using self-organization, including self-multiplication and self-extinction and (2) by collating the input images with the data in the database by means of the low-order image characteristics, while narrowing the search space in accordance with the past history.

In the preliminary construction of the database, we generated CG images of the hand by measuring the joint angles using a data glove and interpolating them; furthermore, we

extracted the contours, image characteristics and the characteristics that change only in the hand shape, irrespective of the environmental light or skin colour. The image was divided into several images and was converted into a number of characteristics by using the high-order local autocorrelation function; the image was then saved in the database in a state paired with the joint angle data obtained from a data glove. By clustering this database using self-organization depending on the number of characteristics and by the self-organization of classes in multiple stages, a multistage search was enabled using the representative numbers of classes in several layers. Moreover, by incorporating self-multiplication and self-extinction algorithms, we achieved a unification of the amount of data belonging to each class as well as the number of classes in the lower layers to avoid the dispersion of the search time in the classes.

The input image at the time of an actual estimation of the hand finger shape was subjected to various types of image processing techniques in the same manner as that at the time of construction of the database, and it was converted into a number of characteristics. The distance from the number of characteristics obtained from the picture was calculated by using a representative number of characteristics. Classes at close distances were selected as candidate classes for the estimated angle, and a similar distance calculation was also performed in the classes in each layer belonging to a candidate class for the estimated angle. Among the respective data belonging to the candidate classes for the estimated angle in the lowest class, the angle data of the data with the closest distance between the number of characteristics was considered as the estimation result. Furthermore, for the selection of a candidate class, we attempted to reduce the search space by using the previous estimation results and the neighbour information.

By estimating the sequential images of the finger shape by using this method, we successfully realized a process involving a joint angle estimation error within several degrees, a processing time of 150 - 160 fps, and an operating time without dispersion by using a PC having a CPU clock frequency of 2.8 GHz and a memory capacity of 512 MB. Since the image information and the joint angle information are paired in the database, the system could reproduce the same actions as those of the fingers of a human being by means of a robot without any time delay by outputting the estimation results to the robot hand.

5. Acknowledgement

This work is partly supported by Proposal-Oriented Research Promotion Program (PRESTO) of Japan Science and Technology Agency (JST) and Solution-Oriented Research for Science and Technology (SORST) project of JST.

6. References

Athitos, V. & Scarloff, S. (2002). An appearance-based framework for 3D hand shape classification and camera viewpoint estimation, *Proc. Automatic Face and Gesture Recognition*, pp.40-45

Bernardin, K.; Ogawara, K.; Ikeuchi, K. & Dillmann, R. (2005). A sensor fusion approach for recognizing continuous human grasping sequences using Hidden Markov Models, *IEEE Transactions on Robotics*, Vol.21, No.1, pp.47-57

Gallese, V.; Fadiga, L.; Fogassi, L. & Rizzolatti, G. (1996). Action recognition in the premotor cortex, *Brain*, Vol.119, pp.593-60

Hoshino, K. & Tanimoto, T. (2005). Real time search for similar hand images from database for robotic hand control, *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, Vol.E88-A, No.10, pp.2514-2520

Hoshino, K. & Tanimoto, T. (2006). Method for driving robot, *United Kingdom Patent Application No.0611135.5*, (PCT/JP2004/016968)

Hoshino, K. & Kawabuchi, I. (2005). Pinching at finger tips for humanoid robot hand, *Journal of Robotics and Mechatronics*, Vol.17, No.6, pp.655-663

Hoshino, K. & Kawabuchi, I. (2006). Hobot hand, *U.S.A. Patent Application No.10/599510*, (PCT/JP2005/6403)

Kameda, Y. & Minoh, M. (1996). A human motion estimation method using 3-successive video frames, *Proc. Virtual Systems and Multimedia*, pp.135-140

Kohonen, T. (1988). The neural phonetic typewriter, *IEEE computer*, Vol.21, No.3, pp.11-22

Lu, S.; Metaxas, D.; Samaras, D. & Oliensis, J. (2003). Using multiple cues for hand tracking and model refinement, *Proc. CVPR2003*, Vol.2, pp.443-450

Otsu, N. & Kurita, T. (1998). A new scheme for practical, flexible and intelligent vision systems, *Proc. IAPR. Workshop on Computer Vision*, pp.431-435

Rehg, J. M. & Kanade, T. (1994). Visual tracking of high DOF articulated structures: an application to human hand tracking, *Proc. European Conf. Computer Vision*, pp.35-46

Rizzolatti, G.; Fadiga, L.; Gallese, V. & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions, *Cognitive Brain Research*, Vol.3, pp.131-141

Appendix: Composition of the humanoid robot hand

(Hoshino & Kawabuchi, 2005; 2006)

As compared to walking, the degree of freedom (DOF) assigned to manipulation functions and to fingers is extremely low. The functions of the hands are mostly limited to grasping and holding an object and pushing a lever up and down. The robot hand itself would tend to become larger and heavier and it would be almost impossible to design a slender and light-weight robot if currently available motors and reduction gears are used with a number of DOFs equivalent to that of the human hand. It is important to determine where and how to implement the minimum number of DOFs in a robot hand.

We have designed the first prototype of a dexterous robot hand. The length from the fingertip to the wrist is approximately 185 mm and the mass of the device is 430 g, which includes mechanical elements such as motors with encoders and reduction gears without electrical instrumentation such as motor control amplifiers, additive sensors, or cables for external connection.

Fig.13 shows two examples of generating movements involved in Japanese sign language. In the case of the numeral 2, the index finger and the middle finger should be stretched during abduction and pass through a clearance generated by the thumb. Generating the numeral 30 involves a difficulty. A ring is formed by the thumb and the fourth finger and the other three fingers are stretched while exhibiting abduction and then bent to a suitable angle. As for the two examples generated by this system, movements were carried out promptly while maintaining an appropriate accuracy in order to facilitate a reasonable judgment of the numerals created by using the sign language. The time duration of the movement is slightly over 1 s for the numeral 2 and approximately 2 s for the numeral 30.

An important function for the robot hand is picking up small, thin, or fragile items using only the fingertips. This capability is equally or even more important than the ability to

securely grasp a heavy object. Therefore, we designed a second prototype focusing on the terminal joint of the fingers and the structure of the thumb.



Fig. 13. Examples of the sign language movements.

As an independent DOF, we implemented a small motor at every fingertip joint, namely at distal interphalangeal (DIP) joints of four fingers and interphalangeal (IP) joint of the thumb. The mass of the motor is approximately 10 g with a gear. Although the maximum motor torque is very small (0.5 Nmm), the maximum fingertip force is 2 N because of the high-speed reduction ratio and the short distance between the joint and the fingertip, which provides sufficient force for picking up an object. Moreover, it has a wide movable range. Each fingertip joint can bend inward as well as outward, which, for instance, enables the robot hand to stably pick up a business card on a desk.

We also added a twisting mechanism to the thumb. When the tips of the thumb and fingers touch, the contact is at the fingertip and the thumb pads; however, this may not provide a sufficient contact with the other fingertip pads since the thumb cannot twist to make this contact. The human hand has soft skin and padding at the fingertips and the high control of motion and force at the fingertips enables stable pinching even if the finger pads are not in complete mutual contact. However, we expect that the fingertip force produced by the terminal joint drive at the tip of the two finger groups will act in opposite directions at the same point, implying that the two fingertips will oppose each other exactly at the pad.

Fig.14 shows the snapshots of the performance of the second type of robot hand, which repeated the series of movements and stably pinched the business card. The mass of the hand is approximately 500 g and the length from the fingertip to the wrist is approximately 185 mm, which are almost equivalent to those of the human hand.



Fig. 14. Snapshots of the robot hand handling a business card using two or three fingers.



Humanoid Robots: New Developments

Edited by Armando Carlos de Pina Filho

ISBN 978-3-902613-00-4

Hard cover, 582 pages

Publisher I-Tech Education and Publishing

Published online 01, June, 2007

Published in print edition June, 2007

For many years, the human being has been trying, in all ways, to recreate the complex mechanisms that form the human body. Such task is extremely complicated and the results are not totally satisfactory. However, with increasing technological advances based on theoretical and experimental researches, man gets, in a way, to copy or to imitate some systems of the human body. These researches not only intended to create humanoid robots, great part of them constituting autonomous systems, but also, in some way, to offer a higher knowledge of the systems that form the human body, objectifying possible applications in the technology of rehabilitation of human beings, gathering in a whole studies related not only to Robotics, but also to Biomechanics, Biomimmetics, Cybernetics, among other areas. This book presents a series of researches inspired by this ideal, carried through by various researchers worldwide, looking for to analyze and to discuss diverse subjects related to humanoid robots. The presented contributions explore aspects about robotic hands, learning, language, vision and locomotion.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Kiyoshi Hoshino (2007). Copycat Hand - Robot Hand Generating Imitative Behaviour at High Speed and with High Accuracy, Humanoid Robots: New Developments, Armando Carlos de Pina Filho (Ed.), ISBN: 978-3-902613-00-4, InTech, Available from: http://www.intechopen.com/books/humanoid_robots_new_developments/copycat_hand_-_robot_hand_generating_imitative_behaviour_at_high_speed_and_with_high_accuracy

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2007 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen