The Application of Dynamic Analysis
to Hand Gestures

Toshiya Naka[1,2(✉)]

[1] Kyoto University, Yoshida Honmachi Sakyo-ku, Kyoto 608-8501, Japan
[2] Panasonic Corporation, Kadoma City, Osaka 571-8501, Japan
naka.tosiya@jp.panasonic.com

Abstract. In designing new user interfaces for the latest mobile devices, such as tablets and smartphones, intuitiveness and simplicity are very important factors. Interacting with hand-gestures is one of the best choices for enabling the ease of use of such products. However, there are several disadvantages to this kind of intuitive interface: one of the chief problems with gesture-based interaction is that it is difficult to distinguish reliably between unconscious and intentional gestures. To resolve this ambiguity, the authors have proposed a quantitative analysis method of human gestures using dynamics. We discovered a close correlation between intended gestures and the torque applied to each joint; however, our previous model was designed for the quantitative analysis of gesture interaction mechanisms in full-body motions. In this paper, we expand our dynamic motion analysis to finger gestures, and reveal that our proposed method is applicable to dynamic motion analysis of basic finger operations.

Keywords: Hand · Gesture · Dynamics · Easy-to-use · Human interaction

1 Introduction

In human communication, it is well known that gestures can be a richer channel of communication than language; we frequently use hand-gestures unconsciously, such as waving to say goodbye or beckoning with the hand. In designing new user interfaces for the latest mobile devices, interaction with hand-gestures is widely adopted for reasons of ease of use. The latest hand-gesture recognition technologies include *PrimeSense* (2010), which uses an infrared projector, camera and a special microchip to track the movement of our body in three dimensions [1]. The *Microsoft Kinect* (2015) adopted this technology in the *XBox* with gesture operation [2]. *Leap Motion* (2015) is a device specifically targeted for hand-gesture recognition that provides a limited set of relevant points [3], and *Google's Soli* (2015) is a major new gesture technology which uses miniature radar with high positional accuracy to pick up slight movements without the need to touch the device itself [4]. Unlike full-body gestures analysis, accurate motion tracking systems will be required to measure finger movement precisely. The most recently developed technology in this field, *Perception Neuron* (2014), employs a capture system which uses up to 32 inertial measurement units (*IMU*s) to track full-body motions. These *IMU*s have a gyroscope, accelerometer, and magnetometer, and can measure finger movements simply and accurately [5]. Using these latest

motion capture devices has made it possible to capture and analyze human gestures relatively inexpensively.

On the other hand, a great deal of research has been carried out in the field of hand gesture analysis. Rautaray and Agrawal (2015) briefly discussed the use of hand gestures as a natural interface for gesture taxonomies, their representations and recognition techniques, software platforms and frameworks [6]. Panwar (2012) presented a real-time system for hand gesture recognition that relies on the detection of meaningful shapes, based on features like orientation, center of mass, status of fingers and thumb in terms of raised or folded fingers, and their respective locations in images [7]. Meng et al. (2012) have proposed a new approach to hand gesture recognition, which is accomplished by dominant points-based hand finger counting using skin color extraction [8], and Dominguez et al. (2006) suggested a unique vision-based robust finger tracking algorithm in which they used to segment out objects by encircling them with the user's pointing fingertip to be robust to changes in the environment and user's movements [9]. In the new perspective of researchers' studies, the authors tried to clarify the dynamic mechanisms of certain characteristic behaviors, and revealed that some special gestures were quantified by the torque values of elements of the human skeletal model (Naka et al. 2016 [10]). Their basic idea was that human tend to apply greater forces than normal to the relevant portion of our arms or body to emphasize a particular action; it is therefore possible to quantify the dynamic effects in terms of the torque applied to each joint. By selecting hundreds of characteristic gestures and applying them to the proposed model, the authors found that it is possible to represent the degree of exaggeration in a quantitative manner, and found out that their model was applicable to the speaker's emphasized movements for attracting the audience's attention during speeches or presentations, too (rhetorical emphasis). There was a close correlation between the intended gesture and the applied torque.

In this paper, we will expand our proposed dynamic gesture analysis model to finger gestures by defining the hierarchical structure of the hand. Generally, there are structural differences such that the *DOF* (degree of freedom) of each joint is one or two, with the exception of the thumb joint. The total values derived from dynamic analysis of the fingers are much smaller than for the body, so to address these problems, the authors had to analyze the effects of twisting torques more precisely, and needed to consider how to improve the *SNR* (signal-to-noise ratio) for smaller torque values. We will describe our dynamic gesture analysis of finger movement in detail in the following sections.

2 Hand Gesture Analysis Model

In this section, a basic dynamic model and algorithm will be defined and verified to be able to accurately analyze finger gestures. In general, finger gestures can be expressed in the form of a hierarchical structure. Each parameter of the link model is shown in Fig. 1. The human body is typically built as a series of nested joints, each of which may have a link associated with it and facing in the +z direction with +y up and +x to the left (b). In the following experiments, we set the first target operation with fingers to the most popular touch panel (a). The latest touch devices are usually equipped with a

mechanism for detecting the pressure of the fingers. It is also one of the most suitable applications for analyzing these mechanisms by using dynamic analysis (e.g., the correlation between pressure and torque). In Fig. 1, *DIP*, *PIP* and *MP* represent the distal interphalangeal, proximal interphalangeal and metacarpophalangeal joint, respectively.
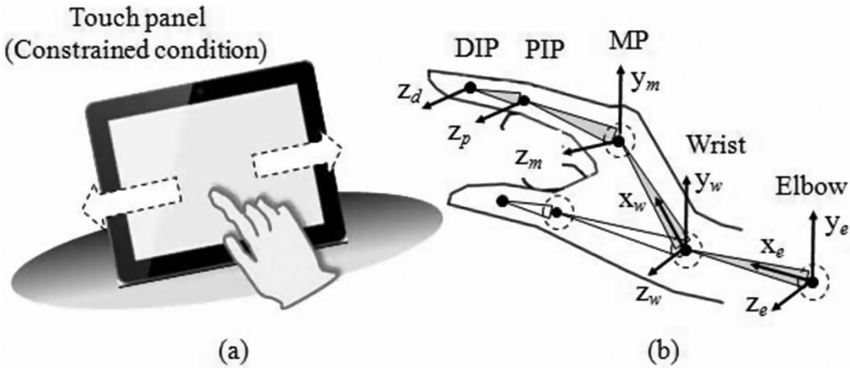

Fig. 1. (a) Finger gesture operation of touch panel (constrained conditions). (b) Hierarchical structure and definition of right hand. The reference (root) position in the dynamic analysis is set to the position of the elbow joints.

Once the structure of the human finger is defined using this hierarchical structure, any finger gestures can be quantitatively expressed as the rotational angles of the time-series around the x, y and z axes (local coordinate system) of each joint, such as the *DIP*, *PIP* and *MP*, and dynamical torque τ which is generated at each joint can be obtained using motion Eq. 1, employing joint angle θ. In this equation, θ is each joint's rotational angle in a time-series data set. $(\theta_w, \theta_m, \cdots \theta_d)$, M is the inertia matrix, C is the *Coriolis force*, g is the *gravity* term and $d\theta/dt$ and $d^2\theta/dt^2$ respectively represent the angular velocity and angular acceleration of each joint. See Naka et al. (2016) for more details [10].

$$\tau = M(\theta)\frac{d^2\theta}{dt^2} + C(\theta, \frac{d\theta}{dt})\frac{d\theta}{dt} + g(\theta) \tag{1}$$

As previously mentioned, it should be noted that in the dynamic analysis of finger movements, these values are noisy besides very small compared with the magnitudes of body values. Highly accurate measurement of angular change θ, noise removal and more precise motion prediction is therefore the key to these analyses.

3   Experiments and Results

To investigate what degree of accuracy is necessary in the analysis of finger move-
ments, we conducted some preliminary experiments. The authors first measured the
dynamical torque of each finger while operating the touch panel shown in Fig. 1(a). In
this series of operational tasks, users usually employ only the index finger, and will
move the *DIP*, *PIP*, *MP* and wrist joints only as required. Normally, this operation is
carried out on the two-dimensionally constrained surface of a touch panel. The authors
term this finger gesture operation as under "constrained conditions."

3.1   Experimental Conditions

In the following experiments, a data glove was used for motion tracking to measure
finger gestures precisely. The main specifications of this system were listed in Table 1.
Subjects were instructed to wear a data glove on the dominant hand and each motion
was converted to each joint rotational angle θ in time-series data (60 fps). The latency
of calculation was order of 10 to 20 ms and the data was translated from Hub to
Computer by using wired USB (in a few ms). The conversion from angle θ to torque τ
was executed about 5 ms on PC by using Eq. 1 (See Appendix for more detail
sequences [10]). In general, each finger gestures θ is noisy and the total values derived
from dynamic analysis of the fingers are much smaller than for the body, so to address
these problems, we had to remove the noise by the low pass filter and adaptive cutoff
frequency of the filter was selected a hundred to two hundred Hz to improve the *SNR*
(signal-to-noise ratio) for smaller torque values. As for the motion prediction, we used
the *three dimensional spline function* to estimate to track smooth the motion of finger
gestures.

Table 1.   Main specifications of data glove

Main specifications	
Number of sensors	5 to 18 (Repeatability: 3°) two sensors per finger
Sensor resolution	<1.0°
Sensor data rate	30/60 fps
Sensor linearity	0.6% maximum nonlinearity over full range

Twelve adults in their twenties (nine men and three women) were selected as
subjects, and each subject was instructed to manipulate the graphical user interface
(*GUI*) by using only their index finger gestures on the constrained touch panel. The
main tasks of these basic experiments were the simple operation of scrolling a page up
and down or to the left and right with the index finger.

3.2   Results 1

Figure 2 shows some typical analysis results of torque values of finger gesture operations. In this figure, (a) shows the operation of turning a GUI page left and (b) upwards with their index finger. In these results, τ_{mp-y} is the joint torque of the MP joint around the y-axis, τ_{mp-z} is the torque of the MP around the z-axis and τ_{mp-x} shows the torque of the wrist joint around the x-axis. The results of a series of preliminary experiments such as these showed that the required accuracy for analyzing finger operation could be obtained in the system environment shown in Table 1. However, there was also a need to create certain prediction methods and apply low-pass filtering to remove noise during each motion (Dominguez et al., 2006 [9]). With these tasks under "constrained" conditions, it should be possible to operate a GUI in parallel too, but slightly distant from the touch panel surface. The dynamic analysis results under "constrained" conditions showed only a small value for τ_{mp-x} and τ_{mp-z} the change in twisting torque around the wrist joint, as shown in Fig. 2.
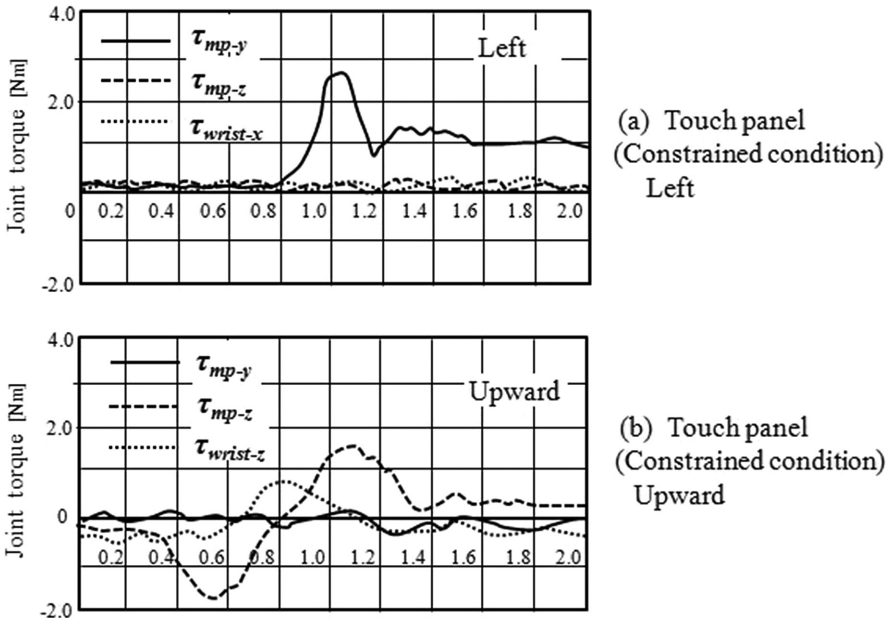

Fig. 2. Typical experimental results for torque values of finger gestures: (a) turning a page to the left and (b) upwards with the index finger. In this figures, τ_{pip-z} is the joint torque of the PIP around the z-axis, both τ_{mp-y} and τ_{mp-z} are the torque of the MP around the y-axis and z-axis, $\tau_{wrist-z}$ is the torque of the wrist around the z-axis, also $\tau_{wrist-x}$ and $\tau_{wrist-z}$ show the torque of the wrist joint around the x-axis and z-axis.

In the next experiment, we attempted to apply dynamic analysis to another typical finger gesture, which in this case was the natural and unconstrained motion shown in Fig. 3. Users often want to control displays with gestures at a distance from them,

particularly if not able to touch the display directly due to having wet or dirty hands. These "constraint-free operations" are frequently reported as feeling natural, but in fact they tend to be difficult for inexperienced users because of too many degrees of freedom. To verify these facts mathematically, we attempted to analyze these types of finger gestures dynamically.
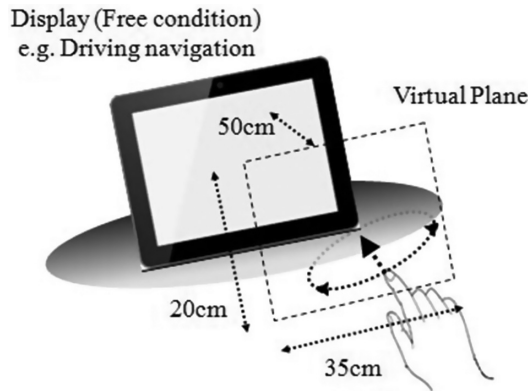

Fig. 3. Typical finger gesture operation under "free" conditions e.g. driving navigation. Usually this is necessary when operating a display using gestures at a distance from the screen. "Virtual plane" placed a transparent plane that was thin and transparent plastic plate (20 cm in length and 35 cm in width)" in 50 cm front of the display

3.3 Result 2

Figure 4 shows typical torque values in the time domain of finger gestures. In these experimental results, τ_{mp-y} is the joint torque of the MP around the y-axis, τ_{mp-z} is the torque of the MP around the z-axis, τ_{pip-z} is the torque of the PIP around the z-axis, and $\tau_{wrist-x}$ shows the torque of the wrist joint around the x-axis. All the subjects noted this to be more difficult than under the "constrained" conditions shown in Fig. 1. Qualitatively, rotational movement around the wrist joint was dominant due to the wrist position as the base not being fixed in this unconstrained situation. Figure 4 shows the typical dynamic analysis results describing such feelings. A comparison of the results in Figs. 2 and 4 suggest that the twisting torques of the wrist joint were observed to dominate during "free condition" operation. Most subjects usually operated the *GUI* using the *MP* joint around the y-axis for (a) moving left and (b) around the z-axis for moving upwards. It appears from the results of analysis that the higher the values of the torque around the wrist joint, the more unstable the operations tend to become.
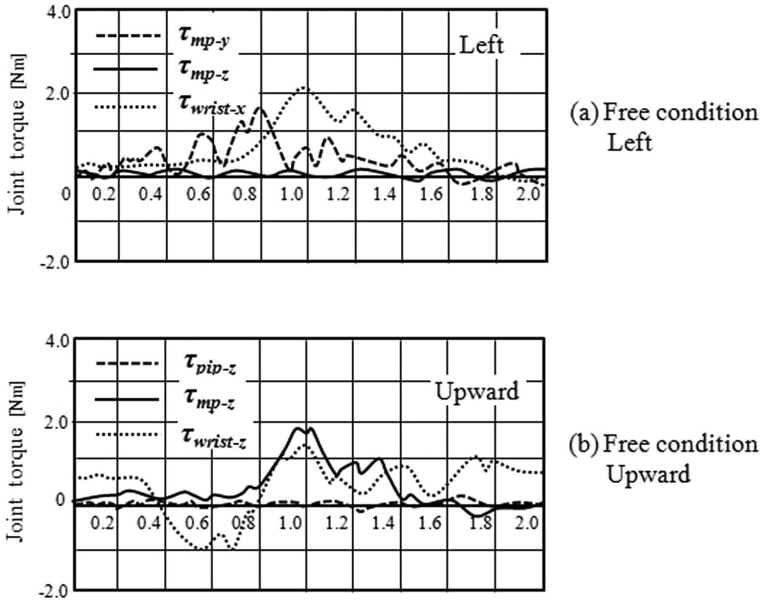
Fig. 4. Typical experimental results of torque value of finger gestures: (a) turning a page to the left. (b) upwards with the index finger ("free" condition). In these figures, $\tau_{pip\text{-}z}$ is the joint torque of the PIP around the z-axis; both $\tau_{mp\text{-}y}$ and $\tau_{mp\text{-}z}$ are the torque of the MP around the y-axis and z-axis, $\tau_{wrist\text{-}z}$ is the torque of the wrist around the z-axis, also $\tau_{wrist\text{-}x}$ and $\tau_{wrist\text{-}z}$ show the torque of the wrist joint around the x-axis and z-axis respectively.

4   Hypothesis and Verification

The authors propose the following hypothesis as a method of overcoming the difficulty faced due to completely free operation with no restraints: they tried to imagine the existence of a restraining surface above the actual surface, as shown in Fig. 3. Under these "pseudo-constrained" conditions, the task feels easier due to a perceived reduction in DOF. As a verification experiment, we placed a transparent plane that we term a "virtual plane (Thin and transparent plastic plate 20 cm in length and 35 cm in width)" in front of the display and asked the subjects to complete the same tasks as in Fig. 1. When asked about how easy they felt the task was, the same twelve subjects answered that it was considerably improved. The results of careful dynamic analysis of the finger gestures during these operations showed that almost all the results were uniformly very similar to the torque changes under the restrained conditions shown in Fig. 2. In other word, the torque values around their wrists had been uniformly suppressed.

4.1   Proposal for a New Improved Ease of Operability

In a completely free operating as shown in Fig. 3, there is potential to improve ease of use by suppressing the torque of the wrist joint around the x-axis. In general, to

improve the reliability of the gesture operation, it appears that some tactile feedback would be effective. Electrical stimulation and air pressure have been proposed as a way of providing non-contact tactile feedback (Hachisu et al., 2014 [11]). With these mechanisms, it is likely that tactile feedback will improve if its effect is to counteract the change in torque curve in time series around the wrist $\tau_{wrist-x}$ as shown Fig. 4(a).

5   Conclusion

In this paper, the authors proposed an expanding dynamic motion analysis method for finger gestures. There are some structural differences with the original proposition that was designed for the whole body, such as the degree of freedom (DOF) of each joint, and the cumulative values of dynamic analysis of fingers are much smaller than for the whole body (Naka et al. 2016 [10]). To address these problems, we constructed a high-accuracy measurement system for finger movements and a noise removal method. As the first step, we focused on finger operation of touch panels, which are widely used in mobile phones and tablets, and compared the dynamic mechanisms of a basic gesture under both constrained and free conditions. We obtained the following results from the series of experiments carried out to verify the mechanism quantitatively.

1. The required accuracy for analyzing finger operations could be guaranteed by the system environment shown in Table 1. However, several prediction methods and a low-pass filtering process to remove noise from each motion were needed. With the tasks under "constrained" conditions, we would operate GUI with parallel direction on the touch panel surface, and the dynamic analysis results under "constrained" conditions showed only a small value for $\tau_{wrist-x}$, the change in twisting torque around the wrist joint, as shown in Fig. 2.

2. We also attempted to apply dynamic analysis to another typical finger gesture, this time without constraint, as shown in Fig. 3. These constraint-free operations were usually reported as feeling natural, but in fact that they were often difficult for inexperienced users due to their being too many degrees of freedom. A comparison of these dynamical experimental results showed that twisting torques around the wrist joint tended to dominate in "free condition" operations. Most subjects usually operated the GUI by using the MP joint around the y-axis to indicate movement to the left (a) and around the z-axis for upward movement (b), so it appears that the higher the values of the torque around the wrist joint, the less reliable the operations were.

3. The authors propose the following hypothesis as a method of overcoming the difficulty associated with completely free operation with no restraints: we attempted to intentionally place a restraint surface in the space as shown in Fig. 3. Under these "pseudo-constrained" conditions, the operation felt a good deal easier because of fewer DOF. We placed a transparent plane called a "virtual plane" in front of the display and asked the subjects to engage in the same tasks on it. Dynamic analysis of these tasks showed almost all the measurements to indicate very similar torque changes to those under restrained conditions. In other words, the torque values around their wrists was uniformly suppressed. To improve the difficulty of

completely free operation with no restraints, it would be suggested to cancel the twisting torque curve change around the wrist $\tau_{wrist-x}$ by using tactile feedback.

The experiments shown in this paper indicate that this approach can be effectively adapted to several basic finger gestures. In future studies, it will be necessary to verify further potential for improvement of this model in terms of accuracy or analysis of more complex finger movements. In addition, we would like to work on a method to more accurately capture and analyze finger gestures.

The authors wish to express their special thanks to Panasonic Corporation's PK-project, which supported this research.

Appendix

As mentioned in the Sect. 2, each dynamical torque τ which is generated at each joint can be obtained using motion Eq. 1, employing joint angle θ which is each joint's rotational angle in a time-series data set $(\theta_w, \theta_m, \cdots \theta_d)$. Since this equation is the general formula of the dynamic model, we have to calculate the torque τ using *Lagrange functions* in the following steps.

1. Step-1: We define the joints' rotational angles by the generalized coordinate system θ_i ($i = 0 \sim$ n) and physical parameters of the approximate finger model as shown Fig. 5.
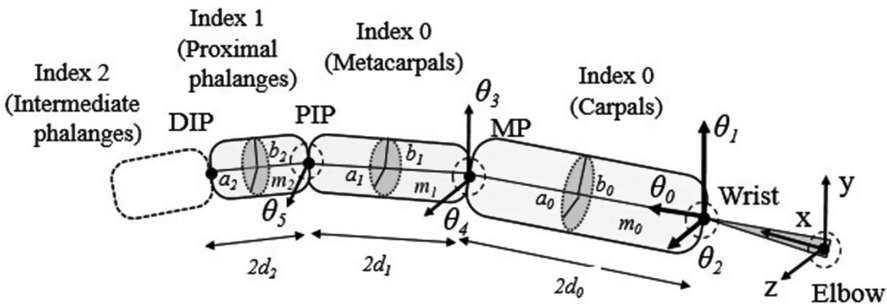

Fig. 5. Generalized finger model for dynamical analysis and definition of joints' rotational angles θi (i = 0 $\sim$ n) and physical parameters. (In case of index finger)

In addition, we used the physical parameter values such as width, length and mass of this approximate finger model as shown in Table 2.

2. Step-2: In generally, *Lagrange function L* of each link structure is given by the following Eq. A.1.

Table 2. The value of each parameter of the approximate model used for calculation: width, length and mass.

	Width [mm]		Length [mm]		Mass [g]
a_0 and b_0	19.4	$2d_0$	100.2	m_0	315.3
a_1 and b_1	15.3	$2d_1$	23.2	m_1	13.5
a_2 and b_2	8.1	$2d_2$	21.4	m_2	2.2

$$L = \sum_{0 \leq i \leq n} \{(\text{Kinetic energy of link } i) - (\text{Potential energy of link } i)\} \qquad \text{(A.1)}$$

3. Step-3: Furthermore *Lagrange equation* of motion Qi is given by the following Eq. A.2.

$$Q_i = \frac{d}{dt}\left(\frac{\partial L}{\partial \theta_i'}\right) - \frac{\partial L}{\partial \theta_i} \quad \text{where i} = 0 \text{ to n} \qquad \text{(A.2)}$$

Moreover, the Equation of motion Qi can be written in the following non-linear ordinary differential Eq. A.3 (the general formula is given by Eq. 1).

$$Q_i = \left[\frac{\partial T_j}{\partial \theta_k} J_j \frac{\partial T_j}{\partial \theta_i}\right] \frac{d^2\theta_k}{dt^2} + \sum_{j=i}^{n}\sum_{k=0}^{j}\sum_{l=0}^{j} trace\left[\frac{\partial 2T_j}{\partial \theta_l \partial \theta_k} J_j \frac{\partial T_j}{\partial \theta_i}\right] \frac{d^2\theta_k}{dt^2}\frac{d^2\theta_l}{dt^2}$$
$$- \sum_{j=i}^{n} m_j g^{\mathrm{T}} \frac{\partial T_j}{\partial \theta_i} S_j$$

$$\text{where i} = 0 \text{ to n} \qquad \text{(A.3)}$$

In this equation, Qi represents the torque in case of rotational motion, the first term of the right side shows the term of angular velocity, the second term is the part of *Centrifugal force* and *Coriolis force* and the third term is part of *gravity*. θ_i is i-th joint's rotational angle and T_j is the coordinate transformation matrix to convert the j-th local coordinate system from the world coordinates. Moreover, J_i is the *Inertia tensor* of the link j, m_i is the mass of i-th link. g^T is the *gravity vector* and S_j denotes the *position vector* of center of mass of link j.

4. Step-4: We approximated by the elliptic cylinder of each link as J_i, the length of the ellipse was $2d_i$, and the widths are a_i and b_i respectively as shown in Fig. 5. Furthermore, it was assumed that the density distribution is constant, and the center of gravity was located at the origin of the length of elliptic cylinder. It has been verified that the generality of dynamic analysis is not lost even if approximated in this way. In this simplified case, *Tensor inertia J_i* of the link i can be approximated by the following Eq. A.4 below.

$$J_i = \begin{bmatrix} d_i^2 m_i/3 & 0 & 0 & d_i/2 \\ 0 & a_i^2 m_i/2 & 0 & 0 \\ 0 & 0 & b_i^2 m_i/2 & 0 \\ d_i/2 & 0 & 0 & 1 \end{bmatrix} \tag{A.4}$$

5. *Quasi-Newton method* was used to execute the calculation on the computer based on the above definition formulae. And the adaptive step width Δt of the iterative calculation was selected to 0.07.

References

1. PrimeSense (2010). http://www2.technologyreview.com/tr50/primesense/, Accessed 24 Oct 2015
2. The Microsoft Kinect (2015) https://dev.windows.com/en-us/kinect, Accessed 24 Oct 2015
3. Leap Motion (2015). https://www.leapmotion.com/, Accessed 24 Oct 2015
4. Soli, "Project Soli" (2015). https://www.google.com/atap/project-soli/, Accessed 24 Oct 2015
5. Perception Neuron, developed the cap (2014). https://www.neuronmocap.com/
6. Rautaray, S., Agrawal, A.: Vision based hand gesture recognition for human computer interaction: a survey. Artif. Intell. Rev. **43**(1), 1–54 (2015)
7. Panwar, M.: Hand gesture recognition based on shape parameters. In: Proceedings of the International Conference on Computing, Communication and Applications (ICCCA 2012), pp. 1–6 (2012)
8. Meng, Z., Pan, J., Tseng, K., Zheng, W.: Dominant points based hand finger counting for recognition under skin color extraction in hand gesture control system. In: Proceedings of the 6th International Conference on Genetic and Evolutionary Computing (ICGEC 2012), pp. 364–367 (2012)
9. Dominguez, M., Keaton, T., Sayed, H.: A robust finger tracking method for multimodal wearable computer interfacing. IEEE Trans. Multimedia **8**(5), 956–972 (2006)
10. Naka, T., Ishida, T.: Dynamic motion analysis of gesture interaction. In: Handbook of Research on Human-Computer Interfaces, Developments and Applications, pp. 35–42. IGI Global (2016)
11. Hachisu, T., Fukumoto, M.: Vacuum touch: attractive force feedback interface for haptic interactive surface using air suction. In: Proceedings of the SIGCHI, pp. 411–420 (2014)