In-hand Manipulation Action Gist Modeling and Generalization from a Data-glove

Gang Cheng, Norman Hendrich, and Jianwei Zhang
Department of Informatics, Group TAMS, University of Hamburg
cheng@informatik.uni-hamburg.de

1. Introduction

In-hand manipulation action gist is a kinetic concept which represents the key finger motions in a manipulation task and widely adapts to different hands. Take a simple example: when we want to grasp an object, the fingers close to touch; this "closing" is the action gist in this scenario. In most cases, we should "close" the fingers but not "open" them for grasping. Guided by the action gist, the object is manipulated into the desired state. The manipulation process is generalized as several compact motion guidelines. On the one hand it becomes easy to remember, on the other it can be easily translated from one entity to another, like the knowledge imparted from the teacher to the student.
This study aims at proposing a feasible in-hand manipulation action gist definition for a robot with an extremely realistic humanoid hand, to enable it to learn in-hand manipulation with a small amount of nonetheless key information.

2. Meta Motion Definition

Meta motion is the basic unit in the in-hand manipulation action gist, it is defined as the finger moving tendency related to the palm, as Fig.1 and Fig.2 show. Five fingers cooperate with the palm to perform in-hand manipulation, each finger has 9 types of meta motion.

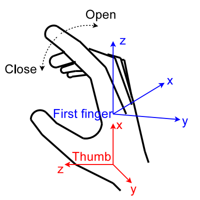


Figure 1: *The finger coordinates with the palm. The thumb in the red coordinate differs from the other four fingers due to its special location in the hand*

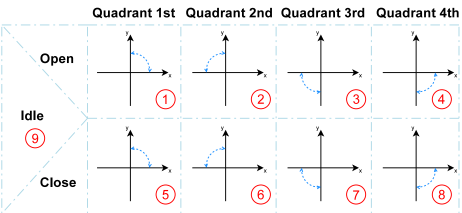


Figure 2: *Nine types of meta motion in each finger. Two flex/ext-joints are modeled as one parameter as open or close, and the abduction angle cooperates with the metacarpal-proximal angle to form a 2D projected direction for the finger's movements. The idle motion is specifically set apart and labeled as 9*

3. Gist Extraction and Generalization

A Gaussian Markov Random Field based algorithm is proposed to extract the action gist of each finger. It can effectively decrease the negative impact from data-glove noise and natural behavior, and finally provide a concise meta motion sequence.

As there are multiple methods to manipulate an object, we have to evaluate the popularity of the action gist sequences in the demonstration set. A Meta Motion Occurrence Histogram is applied to describe the statistical features of the motion order from all demonstration samples.
The work-flow is shown in Fig.3.

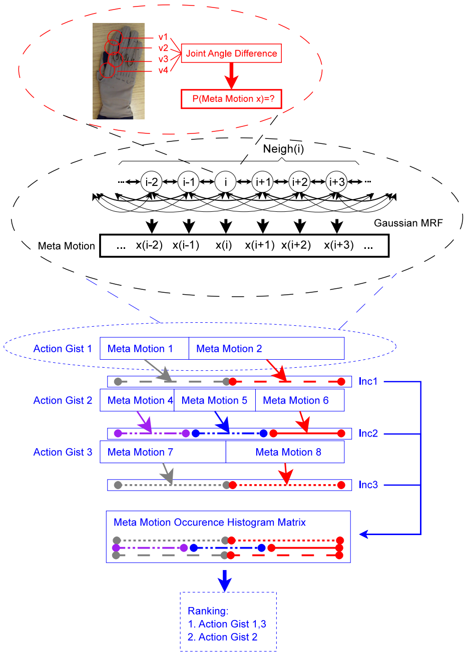
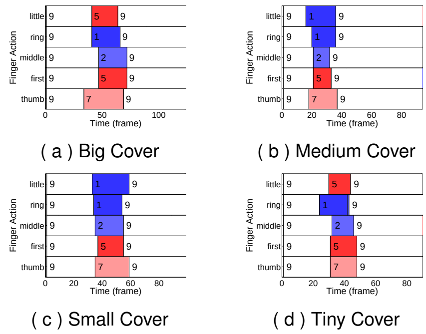


Figure 3: *Workflow of in-hand manipulation action gist extraction and generalization. The value differences between the current frame and the previous frame from the data-glove are used to calculate the meta motion similarity P(Meta Motion x) in each frame. Each single meta motion similarity is considered as a node in Gaussian Markov Random Field; the closer nodes have a stronger influence due to Gaussian distribution. Then the action gist of one trial in the demonstration set is extracted. In order to evaluate multiple action gists, a Meta Motion Occurrence Histogram is applied to count the frequency of each meta motion in every action gist. This gives us a ranking to establish the popularity of the action gists.*

4. Experiment Example

Taking covers opening as an example, we have four different-sized bottle covers as Fig.4. A participant rotate the covers by four fingers for many times. After the trials we can have the action gist result through the proposed method as in Fig.5. The action gist popularity according to each cover in is shown in Tab.1.

Figure 4: *Open different-sized bottle covers. The thumb, first, middle and ring finger participate in this scenario. The cover is rotated as 90 degree anticlockwisely, and this process is defined as a trial*



(a) Big Cover (b) Medium Cover
(c) Small Cover (d) Tiny Cover

Figure 5: *Action gists of covers opening. Each action gist is composed of the meta motions of five fingers, each meta motion is represented by different color rectangles with the corresponding type number. The x-axis is a time axis indicating the cyber-glove frame number. Taking idle motion into consideration, we find that the action gists from 4 trials look similar. The common meta motions are motion 7 in the thumb, 5 in the first finger, 2 in the middle finger, and 1 in the ring finger.*

Table 1: *Action Gist Ranking*

Cover	Action Gist $-\mathbf{m}^s$	Rank
Big	(Thumb,Motion7), (Ring,Motion1), (Middle,Motion2), (First,Motion5)	1
Big	(Thumb,Motion7), (Ring,Motion1), (Middle,Motion1), (Middle,Motion2)	2
Big	(Thumb,Motion7), (Ring,Motion1), (First,Motion5), (Middle,Motion2)	3
Medium	(Ring,Motion1), (Thumb,Motion7), (First,Motion5), (Middle,Motion2)	1
Medium	(Ring,Motion1), (Thumb,Motion7), (First,Motion5), (Middle,Motion2)	2
Medium	(Thumb,Motion7), (Ring,Motion1), (First,Motion5), (Middle,Motion2)	3
Small	(Ring,Motion1), (Thumb,Motion7), (First,Motion5), (Middle,Motion2)	1
Small	(Thumb,Motion7), (Ring,Motion1), (First,Motion5), (Middle,Motion2)	3
Small	(Thumb,Motion7), (Ring,Motion1), (First,Motion6), (Middle,Motion1)	3
Tiny	(Thumb,Motion7), (Ring,Motion2), (First,Motion8), (Middle,Motion2)	1
Tiny	(Thumb,Motion7), (Middle,Motion2), (Ring,Motion4), (First,Motion8)	2
Tiny	(First,Motion4), (Middle,Motion2), (Thumb,Motion7), (Ring,Motion2)	3

5. Conclusion and Future Work

In an action modeling and generalizing form, this study concentrates on the demonstration of in-hand manipulation. Different from manipulator trajectory planning, this model works in a fuzzy way to guide the movement. In the future the model will be examined by simulation and real robot tests.
Our model is currently built from the values of the data-glove. Fusing the results from other sensors is another future direction for development.
The in-hand manipulation process is considered as a *State-Action Model*. This work only discusses the *Action* generation; another important topic is *State* which involves the posture of hand and object, and the contact state. That part is related to visual, haptic and other perceptual channels which are the next goal of our research.

Acknowledgements This work is partially supported by the European FP7 project HANDLE ICT-236410, www.handleproject.eu/.
Gang Cheng acknowledges financial support by the Chinese Scholarship Council.