

# Human-Robot Cooperative Manipulation of Large Deformable Objects

Hideaki Ito  
The University of Tokyo, JSK Lab

## I. INTRODUCTION

In this research, we deal with human-humanoid cooperative-manipulation tasks of large deformable objects where simultaneous operations by two individuals are required. So far, we made a presentation on cooperative operations of large objects with the aim of acquiring the motion of the robot by imitating that of the human [1]. While a motion capture device was attached to the human body to acquire human motion, there were some difficulties remained:

- the sensors can be lost when occlusion occurred during the object manipulation,
- the work space are limited due to the recognizable range of the sensors.

Therefore, in this paper, to acquire the autonomous arm motion of the robot during the task, the target coordinates of the hand are decomposed into two elements: the target position and the target orientation. The former is derived by imitating the human hand movement and the latter by introducing the potential field of the object. Eventually, by combining them, the target coordinates of robot hands are determined. With applying our proposed method, we conducted some experiments with a life-sized humanoid robot HRP-2 and confirmed its effectiveness.

## II. ACQUIRING ROBOT AUTONOMOUS ARM MOTIONS

Fig.1 shows our proposed flow of deciding robot arm motion in cooperative-operation tasks. In this section, first, we will describe the method to acquire the target positions of the robot hands by means of imitation, considering the positions of human hands and face acquired by applying OpenPose [2] (Sec.II-A). Second, we will describe the idea to acquire the target orientation of the robot hands with the consideration of the potential vector field of environmental objects (Sec.II-B).

### A. Acquiring Target Positions of Robot Hands

In situations of cooperative works where a human and a robot manipulate the same object at the same time with facing each other, the target coordinates of the robot hands can be decided by observing human hand motions and imitating them symmetry. Since the acquired three dimensional positions of the human hands are relative to the robot camera, those have to be translated as those relative to the human. Positions of the human hands relative to the human face that are described in the human coordinate system  $\mathbf{p}_{rh}^{hmn}$ ,  $\mathbf{p}_{lh}^{hmn}$  are calculated as below.

$$\mathbf{p}_{rh}^{hmn} = (\mathbf{R}_f^{hum})^T (\mathbf{p}_{rh}^{hmn} - \mathbf{p}_f^{hmn}) \quad (1)$$

$$\mathbf{p}_{lh}^{hmn} = (\mathbf{R}_f^{hum})^T (\mathbf{p}_{lh}^{hmn} - \mathbf{p}_f^{hmn}) \quad (2)$$

$\mathbf{p}_f^{hmn}$ ,  $\mathbf{p}_{rh}^{hmn}$ ,  $\mathbf{p}_{lh}^{hmn}$  refer to the position of the human face and hands relative to the world coordinate system.  $\mathbf{R}_f^{hum}$  refers to a rotation matrix from the world coordinate system to the human coordinate system, which is set at the position of the human face. This is predetermined as a constant matrix below according to the angle between the human and the robot.

$$\mathbf{R}_f^{hum} = \mathbf{R}_z(\theta) = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3)$$

The target positions of the robot right and left hands are decided by referring to the human left and right hand respectively. Target positions of the robot hands relative to the robot face that are described in the robot coordinate system  $\mathbf{p}_{rh}^{rbt}$ ,  $\mathbf{p}_{lh}^{rbt}$  can be acquired as below.

$$\mathbf{p}_{rh}^{rbt} = \mathbf{A} \mathbf{p}_{lh}^{hmn} \quad (4)$$

$$\mathbf{p}_{lh}^{rbt} = \mathbf{A} \mathbf{p}_{rh}^{hmn} \quad (5)$$

$$where \quad \mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Finally, after translating these acquired target hand positions as those defined in the world coordinate system, the target positions of the robot hands  $\mathbf{p}_{rh}^{rbt}$ ,  $\mathbf{p}_{lh}^{rbt}$ , can be acquired, which are the symmetrical imitation of the human hands.

$$\mathbf{p}_{rh}^{rbt} = \mathbf{R}_f^{rbt} \mathbf{p}_{rh}^{rbt} + \mathbf{p}_f^{rbt} \quad (6)$$

$$\mathbf{p}_{lh}^{rbt} = \mathbf{R}_f^{rbt} \mathbf{p}_{lh}^{rbt} + \mathbf{p}_f^{rbt} \quad (7)$$

$\mathbf{p}_f^{rbt}$  refers to the position of the robot head and  $\mathbf{R}_f^{rbt}$  to a rotation matrix from the world coordinate system to the robot face coordinate system, which are predetermined constant vector and matrix.  $\mathbf{R}_f^{rbt}$  is basically equal to the identity matrix  $\mathbf{E}$ ; the orientations of the robot face coordinate system and the world coordinate system are the same.

### B. Acquiring Target Orientations of Robot Hands

Generally, it is hard to recognize the human hand orientation only by a two dimensional image. In order to decide the target orientations of the robot hands, we introduce the idea of potential method. Providing the geometry of the object is known, we introduce the vector field whose direction faces

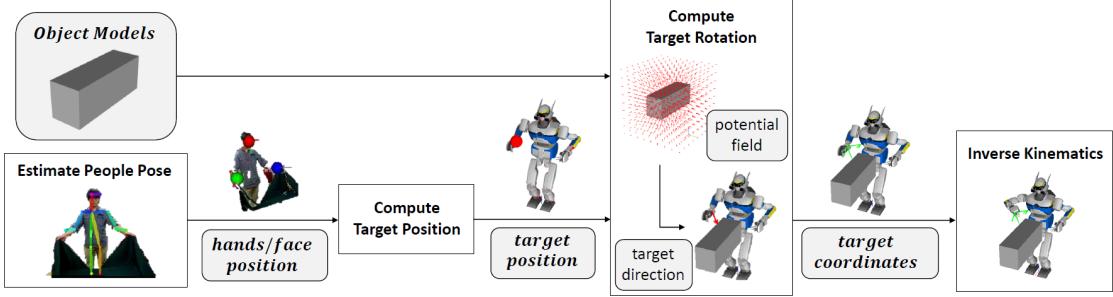


Fig. 1: Overall flow of decision in robot arm motion

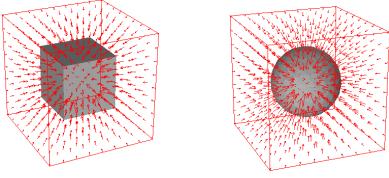


Fig. 2: Examples of vector fields of a cube and a sphere

toward or against the object surface. The target orientation is determined as a posture such that the palm is oriented in that direction. The vector  $\mathbf{E}(\mathbf{r})$  driven by the potential field of the object at a position  $\mathbf{r}$  is calculated as following.

$$\mathbf{E}(\mathbf{r}) = - \sum_{A' \in A} \int_{A'} \frac{\sigma(\mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|^2} \cdot \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|} dA' \quad (8)$$

$A$  refers to the set of the object surfaces, and  $\mathbf{r}'$  to points on the surface  $A' (\in A)$ .  $\sigma(\mathbf{r}')$  refers to the surface density of a virtual charge at the position  $\mathbf{r}'$ , that is, the strength of attraction or repulsion. In this research, this value is set to be constant and to satisfy  $|\sigma(\mathbf{r}')| = 1$ . Note that changing the sign of  $\sigma(\mathbf{r}')$ , the potential field can be switched between attraction and repulsion.

While Eq.8 includes an area fraction, it has to be discretized when simulating. It is done by dividing each surface into small meshes along one set of two adjacent edges and supposing a charge on each meshgrid. Below shows the discretized equation of Eq.8.

$$\mathbf{E}(\mathbf{r}) = - \sum_{A' \in A} \sum_{\mathbf{p} \in MG(A')} \frac{\sigma'}{\|\mathbf{r} - \mathbf{r}'(\mathbf{p})\|^2} \cdot \frac{\mathbf{r} - \mathbf{r}'(\mathbf{p})}{\|\mathbf{r} - \mathbf{r}'(\mathbf{p})\|} \quad (9)$$

Here,  $MG(A')$  is the set of meshgrids of surface  $A'$ . Fig.2 describes the example of vector fields generated by Eq.9 on a cube and a sphere.

The target coordinate of the robot hand is acquired as a posture such that the palm is oriented in the direction at the target position acquired in Sec.II-A.

### III. SEQUENTIAL TASK EXECUTION EXPERIMENT

We conduct three experiments that include manipulations of large deformable objects:

- 1) Wrapping an object with a cloth,
- 2) Putting up a poster on a wall,
- 3) Folding a table cloth.

Fig.3, Fig.4 and Fig.5 show the vector fields of the environmental object and the snapshot in each experiment.

### IV. CONCLUSIONS

In this research, we proposed a method to achieve autonomous arm motion of robot in cooperative manipulation of large flexible objects. In our proposed method, the hand position following human hand movement and hand orientation considering positional relationship with environmental object are combined to derive robot arm motion. The effectiveness was demonstrated through experiments using the real robot HRP-2.



Fig. 3: Wrapping-box experiment



Fig. 4: Putting-up-poster experiment

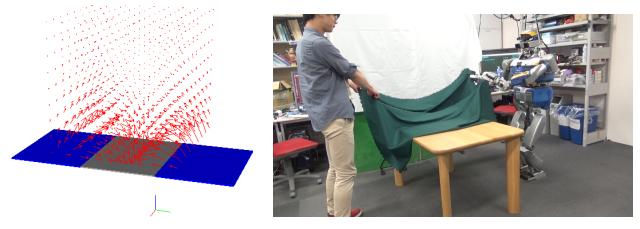


Fig. 5: Folding-table-cloth experiment

### REFERENCES

- [1] 伊藤秀朗, 室岡雅樹, 矢野倉伊織, 野沢峻一, 岡田慧, 稲葉雅幸. 人間の動作模倣と音声指示に基づくヒューマノイドによる対人協調マニピュレーションの実現. 日本機械学会ロボティクス・メカトロニクス講演会'18 講演論文集, pp. 2A2-G12, jun 2018.
- [2] Z. Cao, T. Simon, S. Wei, and Y. Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In *CVPR*, 2017.