Mapping between different kinematic structures without absolute positioning during operation

E. Berglund, B. Iliev, R. Palm, R. Krug, K. Charusta and D. Dimitrov

When creating datasets for modelling of human skills based on training examples from human motion, one can encounter the problem that the kinematics of the robot does not match the human kinematics. Presented is a simple method of bypassing the explicit modelling of the human kinematics based on a variant of the self-organising map (SOM) algorithm. While the literature contains instances of SOM-type algorithms used for dimension reduction, this reported work deals with the inverse problem: dimension increase, as we are going from 4 to 5 degrees of freedom.

Introduction: To train a robot using human demonstrations one needs to measure the position of the human, and translate this into a position for the robot. If there is an absolute reference for the position of the human hand, i.e. a motion capture system, the problem reduces to inverse kinematics (IK). In many cases, motion capture systems cannot be used since they add bulk to the operator's hand and make some motions impossible or unnatural. In these cases one might use an indirect method of measuring the human hand position, e.g. a CyberGlove [1]. We then need to create a transform function to map the CyberGlove data into appropriate commands for the robot.

In these cases it is feasible to split the problem into two: one training phase, where a motion capture system is used, and an operation phase, where a more nimble system is used.

The problem considered in this Letter is distinct from the task of a human controlling the robot directly, as in for example [2], where visual feedback can be used to correct the pose. Other work has been done on programming-by-demonstration [3, 4], whereby a human 'teacher' creates a general pattern the robotic system can emulate, or [5], which uses hidden Markov models to model the hand motion.

In this Letter we apply the parameter-less self-organising map 2 (PLSOM2) [6] algorithm in order to map the input from a human thumb via a CyberGlove sensor glove to a ShadowHand [7] robotic thumb.

Problem description: The problem of mapping the values from the CyberGlove to the ShadowHand is simple in the case of the four ordinary fingers. They have the same number of degrees of freedom (DOF), the joints are of the same type and have the same relative location and orientation. The only remaining unknowns are the lengths of the links and the deflection of the human fingers. Even applying the joint values from the CyberGlove via offset and gain values is enough to get a tolerable accuracy. This is not the case for the thumb, as its structure is fundamentally different from the human thumb. The human thumb has four DOFs: two at the base (combined in a saddle joint) and one each in the medial and distal joints, while the ShadowHand thumb has five, one of which has its rotation axis parallel to the length axis of the thumb.

One way to solve this problem is to use an absolute position system to track the hand and the tip of the thumb in Cartesian space, then solving for the joint angles using inverse kinematics. While this approach will be sufficient to provide accurate joint angles for the robot, it is not always applicable; these systems (e.g. PhaseSpace [8], Polhemus [9]) are bulky and have limited range.

Another solution which would work without an absolute position system would be:

1. Read values from CyberGlove.
2. Estimate human joint angles.
3. Find the position of the human fingertip using forward kinematics.
4. Move the ShadowHand fingertip to the same location using inverse kinematics.

This procedure is outlined in Fig. 1. This approach, which we will call the *naive* approach, relies on having a kinematic model of the human hand in order to estimate the position of the human fingertip. Each of the discrete steps are a potential source of errors. One such approach was described in [3].

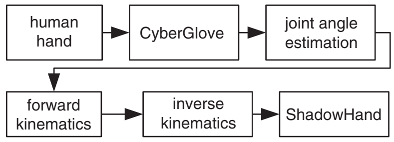

Fig. 1 *Schematic description of a naive approach*

We bypass the indirect modelling by creating a *direct model* that maps from the values returned by the CyberGlove to joint angles for the ShadowHand. Fig. 2 shows how the direct model is created using both Polhemus and CyberGlove data. Fig. 3 shows how the direct model is used to calculate joint angles for the ShadowHand based on data from the CyberGlove. At this point the training is complete, and there is no need for an absolute positioning system.

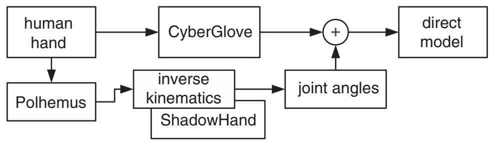

Fig. 2 *Schematic description of learning step of a direct approach*

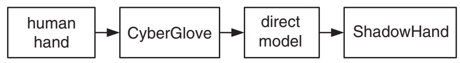

Fig. 3 *Schematic description of operation step*

In the present approach a direct model is obtained by coupling ground truth readings of the fingertip position and orientation from a Polhemus system to corresponding values from the CyberGlove. Based on the position and orientation of the fingertip relative to the rest of the hand, inverse kinematics is used to find joint angles that let the ShadowHand achieve the same position. These joint angles are combined with the CyberGlove readings to create the training set.

Modelling approach: The self-organising map (SOM) [10] family of algorithms (including PLSOM2) can be seen as maps between real spaces of different dimensionality, see (1):

$$f : \mathbb{R}^n \mapsto \mathbb{R}^m \tag{1}$$

where n is the dimensionality of the input space and m is the dimensionality of the output space and $n \geq m$. What characterises the SOM algorithms is that they preserve the topology of the input space [11]. Since there is reason to believe there is an underlying topology (i.e. adjacent joint values are likely to map to similar finger positions), we select the PLSOM2 approach over other possible approaches, e.g. the back-propagation multi-layer perceptron.

In the present case, $n = 4$, the number of values obtained from the CyberGlove for each sample. The output space (joint angles for the ShadowHand) is five-dimensional. To get around this we set the output space to be four-dimensional ($m = 4$) and instead attach a five-dimensional label to each node.

Training the model: Data is collected using a right-handed 26-year-old human male, a 21-DOF CyberGlove, and a Polhemus absolute positioning and orientation sensor system by FCTUC [12]. The CyberGlove is equipped with Polhemus sensors to determine the position and orientation of the palm and of the tip of the thumb. The raw data is available from [13].

The raw data is processed by converting the Polhemus co-ordinates into joint angles for the ShadowHand using IK. The algorithm is tested using five fold cross-validation with 10302 samples in the training set and the remaining 2576 samples in the validation set not seen by the PLSOM2 during training. A PLSOM2 with 4096 nodes arranged in a four-dimensional hypercube and a neighbourhood range of 7 is trained for 100000 iterations. For each iteration one CyberGlove sample is drawn at random from the training set, then applied to the PLSOM2. After training, the map nodes are labelled with their corresponding joint angle values. If more than one joint angle maps to the same node, the centroid is used. Unused nodes are deleted.

The following smoothing algorithm was used to avoid the discrete steps inherent in the model.

We computed an average label weighted by the excitation of the corresponding node, calculated in accordance with (2),

$$z_i = e^{\left(\frac{-d_i - min(-d)}{max(-d) - min(-d)}\right)^{16}} - 1 \tag{2}$$

where z_i is the excitation of node i, and d_i is the distance from the weight vector of node i to the input vector, where the input vector is the normalised data from the CyberGlove. This has the effect of calculating a weighted poll of the best matching map nodes. The excitations are scaled so that (3) is satisfied.

$$\sum_i z_i = 1 \tag{3}$$

Finally the joint values $\vec{v}$ to send to the shadow hand are computed using (4),

$$\vec{v} = \sum_i z_i \vec{l_i} \tag{4}$$

where $\vec{l_i}$ is the label associated with node i.

Results: The validation set is presented to the PLSOM2 one sample at a time. The output $\vec{v}$ is applied to a kinematic model of the ShadowHand thumb, the fingertip position and orientation are compared to the ground truth. The mean distance between the desired location of the fingertip and the actual location is 19 mm, and the mean angle between the desired orientation of the major axis of the fingertip and the actual orientation is 0.47 rad, or about 27°.

Several factors contribute to the error. Each joint monitored by the CyberGlove has a theoretical resolution of 256 discrete steps over the range of the joint. In reality only about half of the steps are actually used, which, at the tip of the thumb, gives an inherent error of circa 2 mm.

While the Polhemus system has a stated RMS error of 1.5 mm, the sensors are mounted on the surface of the CyberGlove and therefore slide relative to the hand as the fingers are moved. Some of the sensors move by as much as 20 mm relative to the hand. Due to their size the centres of the sensors are also offset from the length axis of the fingers they are attached to.

Finally, the carpometacarpal joint of a human thumb allows for up to 3 mm of distraction [14].

For comparison we implemented the naive model using a transpose Jacobian IK solver which was applied iteratively with steps of size 25 nm until convergence. This approach has 44 mm mean error.

The results are described in Table 1, which shows positional error, and Table 2, which shows the deviation of the length axis of the thumb from the target.

Table 1: Position error in millimetres

	Naive	PLSOM2
Mean error	48	19
Median error	47	19
Standard deviation	7	5

Table 2: Orientation error in radians

	Naive	PLSOM2
Mean error	0.83	0.47
Median error	0.81	0.44
Standard deviation	0.08	0.09

Conclusion: By using a direct mapping between the CyberGlove and the ShadowHand we obtain the following advantages:

- Eliminates need for a kinematic model of human hand.
- No need for an absolute positioning system once training is complete.

The proposed method achieves better accuracy than naive methods using CyberGlove data.

Acknowledgment: This work was in part supported by the European Union as part of the HANDLE project.

27 June 2012
doi: 10.1049/el.2012.1085

E. Berglund, B. Iliev, R. Palm, R. Krug, K. Charusta and D. Dimitrov (*Örebro University, Sweden*)

E-mail: erik.berglund@gmail.com

References

1 CyberGlove. [Online]. Available: http://www.cyberglovesystems.com
2 Aleotti, J., and Caselli, S.: 'Interactive teaching of task-oriented robot grasps', *Robot. Auton. Syst.*, 2010, **58**, (5)
3 Palm, R., Iliev, B., and Kadmiry, B.: 'Recognition of human grasps by time-clustering and fuzzy modeling', *Robot. Auton. Syst.*, 2009, **57**, (5), pp. 484–495
4 Ikeuchi, K., Bernardin, K., Ogawara, K., and Dillman, R.: 'A sensor fusion approach for recognizing continuous human grasp sequences using hidden Markov models', *IEEE Trans. Robot.*, 2005, **21**, (1)
5 Asfour, T., Gyarfas, F., and Azad, P.: 'Imitation learning of dual-arm manipulation tasks in humanoid robots'. IEEE-RAS Int. Conf. on Humanoid Robots, (Humanoids 2006), Genoa, Italy, December 2006
6 Berglund, E.: 'Improved PLSOM algorithm', *Appl. Intell.*, 2010, **32**, (1), pp. 122–130
7 Shadow Robot Co. [Online]. Available: http://www.shadowrobot.com
8 PhaseSpace. [Online]. Available: http://www.phasespace.com
9 Polhemus. [Online]. Available: http://www.polhemus.com
10 Kohonen, T.: 'Self-organizing maps' (Springer, 2001, 3rd edn)
11 Kiviluoto, K.: 'Topology preservation in self-organizing maps'. Int. Conf. on Neural Networks, Washington, DC, USA, 1996, pp. 294–299
12 Faria, D.R., Martins, R., Lobo, J., and Dias, J.: 'Extracting data from human manipulation of objects towards improving autonomous robotic grasping', *Robot. Auton. Syst.*, 2012, **60**, (3), pp. 396–410
13 Data used in generating the test data for this paper. [Online]. Available: http://www.aass.oru.se/Research/Learning/data/thumbmovement.zip
14 Brunelli, G.R.: 'Stability of the first carpometacarpal joint' *in* Gilbert, A., Brüser, P. (Eds): 'Finger bone and joint injuries' (Taylor & Francis, 1999), pp. 167–170