Mirror neurons: key for mental simulation?

Erhan Oztop[1], Daniel Wolpert[2], Mitsuo Kawato[1]

[1] ATR, 2-2-2 Hikaridai , Seika-cho, Soraku-gun, Kyoto 619-0288, Japan

[2] Sobell Dept of Motor Neuroscience, UCL, Queen Square, London, WC1N 3BG, UK

Abstract

We propose and implement a model of primate mental simulation for goal directed action recognition, which utilizes neural circuits initially developed for manual manipulation. The model presents a systems level account of how mirror neurons could be part of a mental simulation circuit. We justify the model connectivity and function using the existing monkey neurophysiology and neuroanatomy data. Then we simulate the model in a *hammer-grasping* task where a simulated actor grasps a virtual hammer with different grasping styles, whereas the observer recognizes the action, via mental simulation, paralleling the mirror neuron activity during action observation.

Keywords: Mental simulation; Mirror neurons; Grasping; Action observation

1 Introduction

Humans visually monitor critical kinematic events for detecting errors in goal directed movement execution [1] including grasping movements that require cortical integration of visual and somatosensory cues for proper grip formation [2]. Although the accumulated neurophysiological data indicate that the parietal cortex is involved in visuomotor aspects of manual manipulative movements [3-5], the interplay between visual feedback and feedforward control in grasping is not clear [6]. In this article, we focus on the brain areas involved in gasping and implement a visual feedback control system that we use to show that a preexisting visual feedback circuit could be used in mental simulation mode allowing an action recognition capability.

The neurons in area AIP (anterior intraparietal area) discharge in response to viewing and/or grasping of three-dimensional objects representing object features relevant for grasping [7, 8]. AIP has strong recurrent connection with area F5 (ventral premotor cortex) [9] that is involved in grasp planning and execution [10], and projects to motoneurons that control finger muscles [11]. The activity of neurons in the primary motor cortex (F1) when compared to premotor activity indicates that the primary motor cortex may be more involved in dynamic aspects of movement [12], executing 'instructions' sent by higher motor centers including premotor regions. Thus, it has been suggested that AIP-F5-F1 circuit is responsible for grasp planning and execution [13-16]. An interesting property of this circuit is that F5 mirror neurons in addition to movement related activity, show response to goal directed movements performed by another monkey or an experimenter reflecting the type of actions (e.g. precision or power grasping) [17, 18]. A mirror neuron discharge is usually independent of the depth and the location of the observed movement, and matches the movement related activity of the neuron [17, 18]. However, to our knowledge, there are no studies relating the *temporal* aspects of mirror neuron activity to the *kinematics* of observed movements. Our model predicts that there must be a strong link between hand kinematics with respect to the goal of movement and the mirror neuron activity. The latter prediction and others that are discussed in Discussion can be tested experimentally.

2 The model

2.1 Visual feedback control of grasping and action recognition with AIP-F5 circuit

Based on the findings reviewed above, we suggest that AIP is involved in monitoring the relation of hand with respect to an object, whereas area F5 is involved in converting the AIP output into motor signals, which are used by primary motor cortex and spinal cord for actual muscle activation. In the model, area F5 implements a control policy (assumed to be learned earlier) to reduce the error represented by area AIP output. In the 'observation mode' AIP-F5 functionality is used in a 'mental simulation mode' to mediate action understanding. Although we accept the view that F5 mirror neurons play a key role in generating a mental simulation of grasping [19], we diverge from the generally suggested idea that their goal is simply retrieving a movement representation that matches the observed movement [17, 18]. Instead, we suggest that mirror neurons are part of a manual visual servo circuit, which are *utilized* for mental simulation to understand others' actions. Then, how is 'mental simulation' possible and why mirror neurons discharge for action observation? The answer lies in *forward models* [20-22] that predict sensory consequences of motor related signals without actual movement. It has already been suggested that mirror neurons could be involved in mental simulation [19], but a computational account of *how* was missing. Here we go one step further and specifically propose that F5 mirror neurons implement a *forward model* predicting the sensory consequences of F5 manual manipulation related motor output.

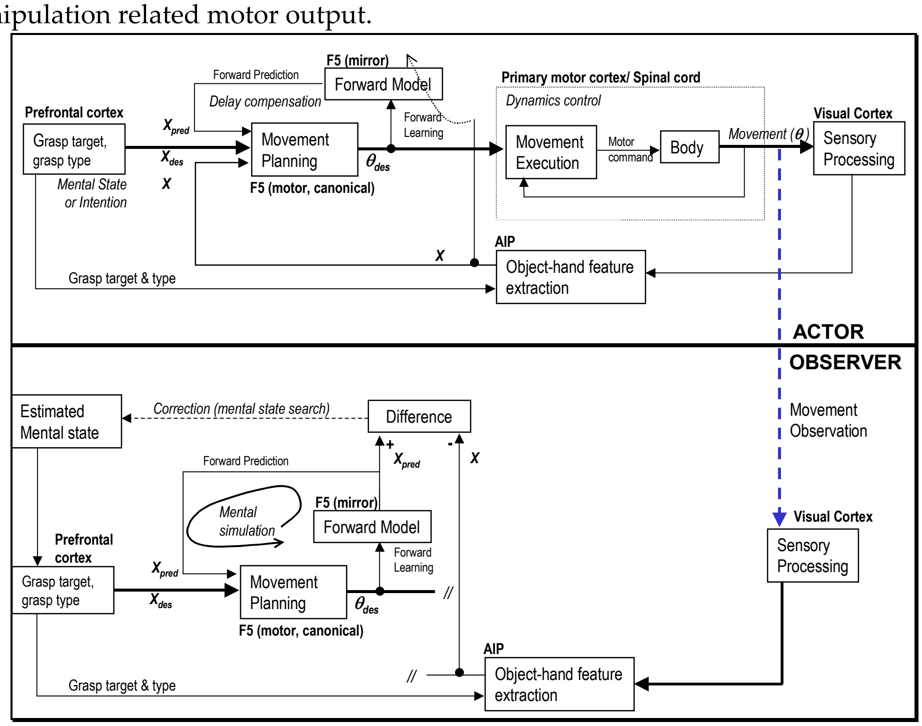

Figure 1: Upper half shows the grasp planning and execution of an actor. Lower half illustrates observer's mental state inference. Mental simulation of movement is mediated by utilizing the sensory prediction from the forward model and by inhibiting motor output. Difference module computes the difference between the visual control parameters of the simulated movement and the observed movement. The mental state estimate indicates the current guess of the observer about the actors mental state (movement goal). In a serial implementation, Difference output is used to update the estimate. In a parallel implementation, the estimate with minimum Difference becomes the observers inference about the actors movement.

The ability of predicting (visual) sensory consequences of movement allows mental simulation as follows. An observer monkey may 'guess' the *target object* (one of possible objects in demonstrator's workspace) and the *type of grasp* and produce an appropriate F5 motor signal that is inhibited for actual muscle activation but used by the forward model (mirror neurons). With the sensory outcome predicted by the mirror neurons, the movement can be simulated as if it were executed in an online feedback mode. The match of the simulated sensations of a simulated movement with the sensation of observed movement will then signal the correctness of the guess (Figure 1). The simulated mental sensations and actual perception of movement is compared in a mental state search as elaborated next.

2.2 Mental State Search

If the observer model 'knows' the possible goals in terms of discrete targets then it can perform an exhaustive search in the mental state space. However, if the number of possible targets is infinite or the targets are not visible to the observer then a different search strategy must be applied. The inferred mental state correction requires the AIP output based errors (the Difference box in Figure 1) to be converted into 'mental state' space adjustments. In general, the difference computation involves an integral over movement time since a movement pattern is characterized by the history of postures achieved during a movement, rather than an instant posture of the limb. In simulation, a discrete version of the Difference can be used with a decay parameter γ that discounts the earlier parts of an observed movement as: $D_N = \frac{(1-\gamma)}{(1-\gamma^{N+1})}\sum_{i=0}^{N}(x_{sim}^i - x^i)^T W (x_{sim}^i - x^i)^T \gamma^{N-i}$. Here, x and x_{sim} represents AIP output and forward prediction, and W is a diagonal matrix normalizing different features extracted by AIP. N is pseudo-time marking the end of movement observed so far. In our simple simulation world, since the possible actions are finite we can use the exhaustive search method in the action space. Whether primates implement a similar algorithm in parallel or serial is not known, however without loss of generality we can present the serial search method:

1. Initialization: T: empty trajectory; A: target of attention; EF: endeffector
2. Repeat 3-7 from movement onset to movement end
3. Shift attention through potential targets (pick target, A)
4. Extract the relevant control variables (CV_{actor}) based on the actor and attended target, A and add to trajectory T_{actor}
5. Mentally simulate movement to target A and compute the mental trajectory of control variables (CV_{self}) and store in T_{self}
6. Let D be the discounted difference between T_{actor} and T_{self}
7. If D is smallest so far, set A_{min}=A

The algorithm returns A_{min}, indicating that observer believes A_{min} is the actor's goal. In the remainder of the paper, we introduce the implementation of the proposed model and present simulation experiments using a grasping task where an observer infers the grasp type of an actor's grasping movement using mental simulation.

3 Simulations

We designed a virtual simulation environment that involves two agents (actor and observer) and a hammer that can be grasped with several ways. The goal of the observer was to infer the type of grasping the actor is engaged in. We first introduce the simulation environment then present the simulation experiments and the details of grasp generation.

3.1 The arm/hand model

The arm is modeled to have 3DOF joints at the shoulder and 1DOF joint in the elbow for lower arm extension/flexion movements. The Wrist is modeled to have 3DOFs to account for extension/flexion, pronation/supination, and ulnar and radial deviation movements of the hand. Each finger except the thumb is modeled to have 2DOFs to simulate metacarpophalangeal and distalinterphalangeal joints of human hand. The thumb is model to have total 3DOFs, one for the metacarpophalangeal joint and the remaining two for extension/flexion and ulnar and radial extension movements of the thumb (i.e. for the carpometacarpal joint of the thumb).

A simple visual feedback controller for a power grasping can be formulated using *distance-to-target* and the *orientation difference* of the hand and the target object's main axis. These parameters comprise a minimal set of parameters that can be used to generate movements for grabbing a hammer. In the simulation, we considered two kinds of grasps: *handle grasp* (for using a hammer as a tool) and *metal-head grasp* (for putting it away) as described next.

3.2 Kinematics for handle grasping

We model the grasping as two parallel visual feedback control problems: transporting hand to the proximity of the center of the handle and adapting the orientation of the *knuckle vector* (see Figure 2A) to match the orientation of the handle. The reaching component of the grasping movement is solved using the pseudo inverse method [23]. The orientation of the hand is adjusted by stochastic gradient descent based on the orientation difference feedback. Thus the control variables enabled the generation of grasping movement were (1) the orientation difference (angle) between the knuckle vector and hammer handle and (2) the distance of the hand to the handle center (see Figure 2A). The mental state search was therefore, based on the distance and orientation variables. The algorithm used for simulating grasping kinematics performed various tasks starting from the onset of movement until the completion of a grasp: (1) reaching towards the handle center via a via-point (to avoid collision) (2) flexing hand finger (3) adapting hand orientation (4) enclosing fingers when collision occurs with the hammer.

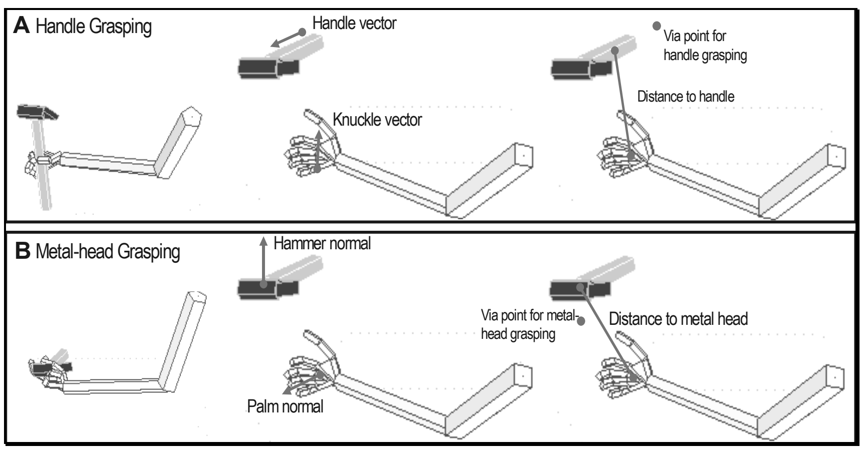

Figure 2: **A**: The features used for handle grasping (orientation and distance) is depicted in the right two arm drawings. The path of the hand is constrained with appropriate via-points avoiding collision. The arm drawing on the left shows an example handle grasp obtained. **B**: The features extracted for metal-head grasping is depicted. (Conventions are the same as the upper panel.)

3.3 Kinematics for metal-head grasping

Grasping of a hammer without a tool-use goal is modeled as power grasping the metal head from a lower approach direction. We model the grasping as two parallel visual feedback control problems: transporting hand to the proximity of the center of the metal head and adapting the hand orientation such that palm normal coincides with the hammer plane normal. Thus, the control variables enabled the generation of grasping movement were (1) the orientation difference (angle) between the palm normal and hammer plane normal and (2) the distance of the hand to the metal head center (see Figure 2B). The mental state search was therefore, based on the distance and normal differences described. The algorithm used for simulating grasping kinematics was the same as handle grasping described above.

4 Results

In general, different actions require different control mechanisms. The hypothesis or the guess of an observer determines which visual features (control variables) have to be computed, even though the actor's movement is generated using different control variables. The success of a mental state search relies on the fact that the minimum of the Difference computation is achieved when the simulated and the actual movement match. We ran two simulations where the actor performed different grasping movements. Since the mental state was small, we recorded observer's belief for each hypothesis separately, effectively implementing a parallel search. In general, an observer can perform the mental state search in parallel or serial. The former is more advantageous for its speed, whereas the latter is more advantageous for its less resource requirement. The output of Difference computation is converted into beliefs using a softmax function, which were computed as follows (*γ=0.9* and *W=diag(0.7,0.3)* normalizing distance and angle units. N represents the simulation cycle, and *x* indicates the parietal and simulated output as *(distance, orientation difference)*^{T}.)

$$D_N = \frac{(1-\gamma)}{(1-\gamma^{N+1})}\sum_{i=0}^{N}(x_{sim}^{i}-x^{i})^{T}W(x_{sim}^{i}-x^{i})^{T}\gamma^{N-i} \quad \textit{and} \quad p(estimate_i \mid movement) = e^{-3D_{Ni}^{i}} \Big/ \sum_{j=1}^{K} e^{-3D_{Ni}^{j}}$$

Figure 3 shows the observer's belief that the actor is engaged in metal-head (left panel) and handle grasping (right hand) while actor is performing a metal-head grasp. Figure 4 shows the belief of the actor in a similar fashion but this time while the actor is performing handle grasping. In both cases, the observer is able to infer the goal of the movement.

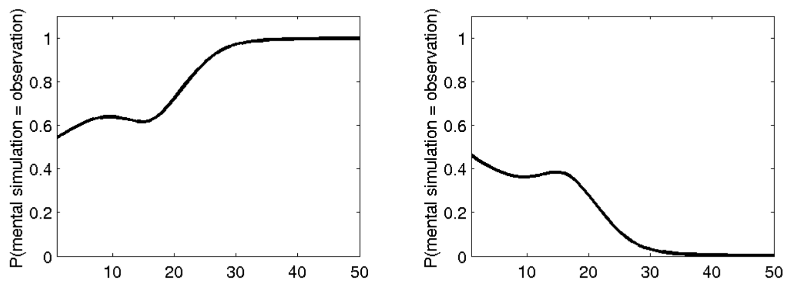

Figure 3: Two plots shows the observer's belief (probability) that the observed movement (metal-head grasp) matches the mentally simulated one (left: metal-head grasping, right: handle grasping).

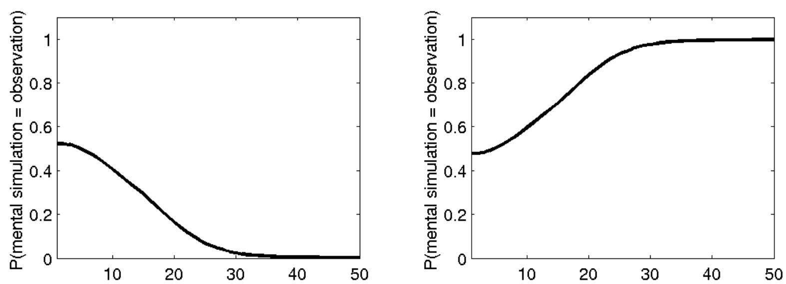

Figure 4: Two plots shows the observer's belief (probability) that the observed movement (handle grasp) matches the mentally simulated one (left: metal-head grasping, right: handle grasping).

5 Discussion

The model presented an explicit account of how the circuits developed for visual feedback control of hand manipulation can be used to understand similar actions performed by conspecifics when the control policy involves a *forward model* that predicts the sensory consequences of movement allowing movement simulation and mental state search. Our model suggests that the mirror neurons found in monkey ventral premotor cortex are involved in *sensory forward prediction* of goal directed hand movements, which are activated for mental simulation during action observation, and for feedback-delay compensation during movement. Thus, we predict that a disruption in mirror neuron activity will *slow down* a grasping movement but will *not* abolish the movement completely when a monkey is tested under conditions inhibiting a feedforward grasping strategy. Feedback control can be ensured by having a target object move or change orientation unpredictably during movement so that the tested monkey have to use a visual feedback strategy for grasping. The predictions of our model can be neurophysiologically verified (or falsified) with lesion studies using a perturbed grasping experiment as described. The last but not the least, our model emphasizes the imperative need for *synchronized* recording of mirror neuron activity with hand kinematics for understanding the mechanisms underlying mirror neuron activity and its cognitive function.

References

1. Johansson, R.S., et al., *Eye-hand coordination in object manipulation.* Journal of Neuroscience, 2001. **21**(17): p. 6917-32.
2. Jeannerod, M., *The formation of finger grip during prehension. A cortically mediated visuomotor pattern.* Behav Brain Res, 1986. **19**(2): p. 99-116.
3. Stein, J.F., *Space and the parietal association areas*, in *Brain and Space*, J. Paillard, Editor. 1991, Oxfort University Press: Oxfort England; New York.
4. Wise, S.P., et al., *Premotor and parietal cortex: corticocortical connectivity and combinatorial computations.* Annu Rev Neurosci, 1997. **20**: p. 25-42.
5. Ferraina, S., et al., *Visual control of hand-reaching movement: activity in parietal area 7m.* European Journal of Neuroscience, 1997. **9**(5): p. 1090-5.
6. Churchill, A., *Vision of the hand and environmental context in human prehension.* Experimental Brain Research, 2000. **134**(1): p. 81-9.
7. Sakata, H., et al., *Neural coding of 3D features of objects for hand action in the parietal cortex of the monkey.* Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 1998. **353**(1373): p. 1363-1373.
8. Sakata, H., et al., *Neural Mechanisms of Visual Guidance of Hand Action in the Parietal Cortex of the Monkey.* Cerebral Cortex, 1995. **5**(5): p. 429-438.
9. Matelli, M., et al. *Independent anatomical circuits for reaching and grasping linking the inferior parietal sulcus and inferior area 6 in macaque monkey*. in *Society for Neuroscience*. 1994.
10. Rizzolatti, G., et al., *Neurons Related to Reaching-Grasping Arm Movements in the Rostral Part of Area-6 (Area-6a-Beta).* Experimental Brain Research, 1990. **82**(2): p. 337-350.
11. Dum, R.P. and P.L. Strick, *The origin of corticospinal projections from the premotor areas in the frontal lobe.* J Neurosci, 1991. **11**(3): p. 667-89.
12. Shen, L.M. and G.E. Alexander, *Preferential representation of instructed target location versus limb trajectory in dorsal premotor area.* Journal of Neurophysiology, 1997. **77**(3): p. 1195-1212.
13. Fogassi, L., et al., *Cortical mechanism for the visual guidance of hand grasping movements in the monkey - A reversible inactivation study.* Brain, 2001. **124**: p. 571-586.
14. Gallese, V., et al., *Deficit of Hand Preshaping after Muscimol Injection in Monkey Parietal Cortex.* Neuroreport, 1994. **5**(12): p. 1525-1529.
15. Jeannerod, M., et al., *Grasping Objects - the Cortical Mechanisms of Visuomotor Transformation.* Trends in Neurosciences, 1995. **18**(7): p. 314-320.
16. Fagg, A.H. and M.A. Arbib, *Modeling parietal-premotor interactions in primate control of grasping.* Neural Networks, 1998. **11**(7-8): p. 1277-1303.
17. Gallese, V., et al., *Action recognition in the premotor cortex*. Brain, 1996. **119**: p. 593-609.
18. Rizzolatti, G., et al., *Premotor cortex and the recognition of motor actions.* Cognitive Brain Research, 1996. **3**(2): p. 131-141.
19. Gallese, V. and A. Goldman, *Mirror neurons and the simulation theory of mind-reading.* Trends in Cognitive Sciences, 1998. **2**(12): p. 493-501.
20. Wolpert, D.M. and M. Kawato, *Multiple paired forward and inverse models for motor control.* Neural Networks, 1998. **11**(7-8): p. 1317-1329.
21. Wolpert, D.M., Z. Ghahramani, and J.R. Flanagan, *Perspectives and problems in motor learning.* Trends in Cognitive Sciences, 2001. **5**(11): p. 487-494.
22. Blakemore, S.J. and J. Decety, *From the perception of action to the understanding of intention.* Nat Rev Neurosci, 2001. **2**(8): p. 561-7.
23. Sciavicco, L. and B. Siciliano, *Modelling and control of robot manipulators*. Advanced textbooks in control and signal processing. 2000, London ; New York: Springer. xxiii, 377.