How humans and artificial classifiers decode grasping movements through kinematic information

Martina Fanghella[1], Fabio Aurelio D'Asaro[2], Maria Teresa Pascarelli[1], Guido Barchiesi[1], Corrado Sinigaglia[1]
[1]Department of Philosophy, University of Milan, Italy
[2]Department of Human Sciences, University of Verona, Italy

Introduction

- Humans use sensorimotor information to understand others' actions (Rizzolatti & Sinigaglia, 2010, 2016), beyond visual recognition of the observed movements (Calvo-Merino et al., 2006)
- Recent studies described the computational basis of action recognition through kinematic information (Cavallo et al., 2016; Montobbio et al., 2022), but no studies manipulated how this varies over time.
- In our study, we used machine learning techniques to describe how kinematic information is used in recognition of grasping movements when variable portions of actions are shown.

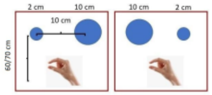
Methods

1. Stimuli acquisition

Two experimenters were asked to reach and grasp two objects (a small ball and a large ball)

160 videos were recorded. Kinematic features (grip aperture and velocity of wrist) were extracted for each video at 10%, 20%, 30% and 40% of movement. Sixteen final videos were selected.

Two 2*2*4 RM ANOVA (size*agent*time) with grip aperture and wrist velocity as dependent variables were performed.



3. Decision tree

Decision Tree was run to test if the algorithm could accurately predict participants' responses on the basis of the kinematic information for each time of movement (10%, 20%, 30%, 40%).

Our dependent variable was participants' responses (small or large) at 10%, 20%, 30% and 40% and our predictors were grip aperture and wrist velocity.

2. Human recognition of grasping actions

Thirty participants (mean age 22.87 ± 0.7; 20 females), all right-handed, were enrolled to the present online study. Participants were asked to discriminate between grasping large or small object by looking at the 10%, 20%, 30% and 40% of the whole movement.

Three 2*4 RM ANOVA (size*time) with recall, precision and reaction times (RT) as dependent variables were computed.

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \qquad Precision = \frac{TruePositive}{TruePositive + FalsePositive}$$

4. Classifier

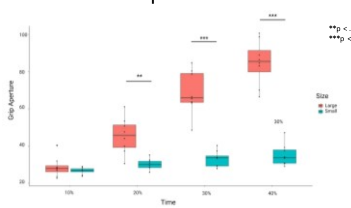
Kinematic data was classified using a classic method known as Support Vector Machine (SVM). We used an RBF kernel with other parameters set as default values.

Visual data was classified using a CNN-RNN Neural Network architecture.

We evaluated our models using 10-folds cross-validation on the 160 videos on recall and precision.

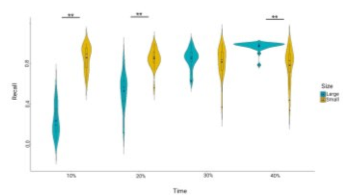
Results

1. Stimuli acquisition



**p < .005
***p < .001

2. Human recognition of grasping actions



3. Decision tree

	10%	20%	30%	40%
Correct responses				
small	100%	89%	89.9%	94.6
large	0%	52%	75.2%	81.3%
overall	81%	76.5%	82.3%	87%
Predictors	a10, a0	a20, a0, a10	a30, a0, a20	a30, v30, a0

4. Classifier

Recall

	10%	20%	30%	40%
Human				
small	0.84	0.83	0.80	0.78
large	0.21	0.50	0.81	0.92
SVM				
small	0.91	0.9	0.98	0.98
large	0.49	0.83	0.88	0.93

Precision

	10%	20%	30%	40%
Human				
small	0.52	0.64	0.84	0.94
large	0.62	0.78	0.82	0.82
SVM				
small	0.66	0.86	0.92	0.94
large	0.74	0.90	0.98	0.98

Discussion

- Decision tree could accurately predict human behaviour using kinematic information (grip aperture).
- SVM classifier using kinematic information to classify grasping actions shows similar patterns of performance compared to human recognition.
- Further analyses with CNN-RNN Neural Network architecture will highlight performance based on visual recognition of movements.

References

- Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or Doing? Influence of Visual and Motor Familiarity in Action Observation. *Current Biology*, *16*(19), 1905–1910. https://doi.org/10.1016/j.cub.2006.07.065
- Cavallo, A., Koul, A., Ansuini, C., Capozzi, F., & Becchio, C. (2016). Decoding intentions from movement kinematics. *Scientific Reports*, *6*(1), 37036. https://doi.org/10.1038/srep37036
- Montobbio, N., Cavallo, A., Albergo, D., Ansuini, C., Battaglia, F., Podda, J., Nobili, L., Panzeri, S., & Becchio, C. (2022). Intersecting kinematic encoding and readout of intention in autism. *Proceedings of the National Academy of Sciences*, *119*(5), e2114648119. https://doi.org/10.1073/pnas.2114648119
- Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. Nature Reviews Neuroscience, 11(4), 264–274. https://doi.org/10.1038/nrn2805
- Rizzolatti, G., & Sinigaglia, C. (2016). The mirror mechanism: A basic principle of brain function. Nature Reviews Neuroscience, 17(12), 757–765. https://doi.org/10.1038/nrn.2016.135