Identifying Kinematic Cues for Action Style Recognition

Shohei Hidaka (shhidaka@jaist.ac.jp)
School of Knowledge Science, Japan Advanced Institute of Science and Technology (JAIST), Japan

Background: Recognition of emotional states from other's actions is one of key capability for smooth social interaction.
Approach: a computational-theory-level analysis kinematic features for recognition of emotional attributes in human actions represented as point-light display.

Results: emotional attributes in could be identified by covariance of velocity profiles distributed among multiple body parts.
Suggestion: the action styles may be mediated by an information channel parallel to action types per se instead of hierarchical manner.

Background

Emotion coordinates social interactions

- To provide information to peers about surrounding environment
- Elicit complementary and similar emotions
- To be an incentive promoting social relationship

Easy to "read" others' attribute through their actions.

- **Identity** (Cutting & Kozlowski, 1977; Troje, Westhoff & Lavrov, 2005)
- **Gender** (Kozlowski & Cutting, 1977; Troje, 2002)
- **Emotions** (Dittrich et al., 1996; Pollick, 2001; Atkinson, 2009; Hobson & Lee, 1999)
- **Dynamics properties** (the weight of lifted object: Bingham, 1987)

But not trivial

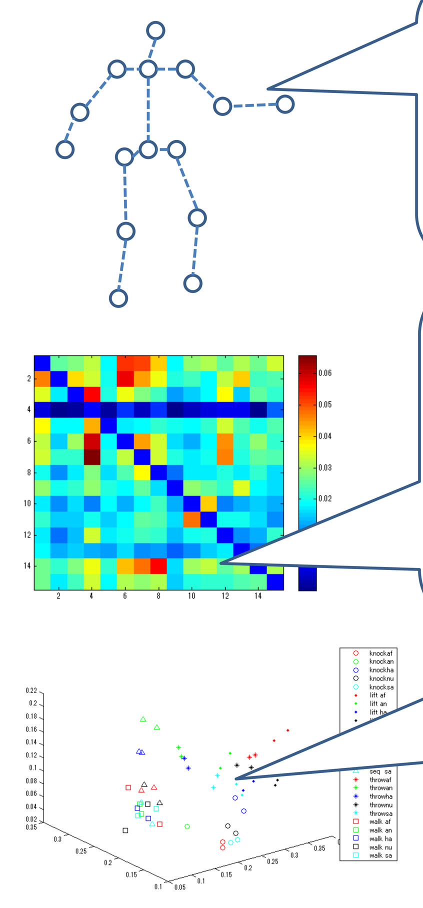
- Different development recognition different types of emotions (Boone & Cunningham, 1998)
- ASD patients can recognize *action types* (lifting, pushing), but *not style* of actions (sad, happy, fear, harsh, gentle) (Moore, et al., 1997; Hubert et al., 2006), can detect coherent random-dot motions, but less with human-biological motion (but evidences are divergent: Kaiser & Shiffrar, 2009).

Model outline

Bio. Motion Library (Ma et al., 2006) 3 actions x 3 emotions

Velocity covariance 15 x 15 (joints)

Classification action/emotion based on info. profile



Subjects: 30 naïve actors with a motion capture (60Hz)
Actions: Lift, knock, throw (15-25 sec)
Emotions: Angry, sad, happy

Velocity covariance profile: Covariance of motion A and motion B.
Markers (joints/ body parts): Head , Neck, Pelvis, (Left/ Right): Shoulder, Elbow, Wrist, Hip, Knee, Ankle

Classification: 3 actions & 3 emotions or 9 action-emotion classes using velocity profiles

Computational processes (hypothesized)

Dynamic cues

- Duration of actions (Pollick et al., 2001)
- Velocity (DeMeijer, 1989)
- Acceleration (2nd order derivatives) (Chang & Troje, 2008; 2009)
- Jerk or (3rd order derivatives) (Cook et al., 2009)

Structural cues

- Body structure (Troje, 2002)

4 levels of recognition of bio. Motion (Troje, 2008)

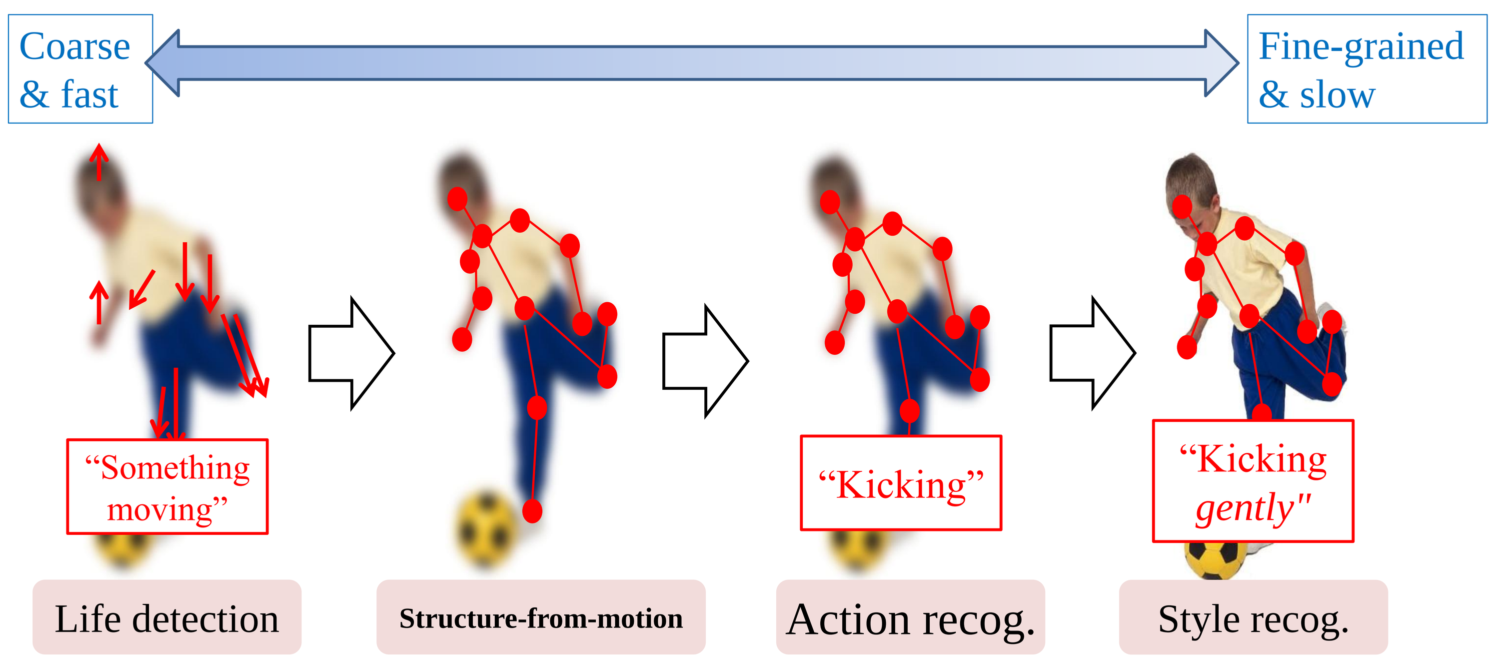
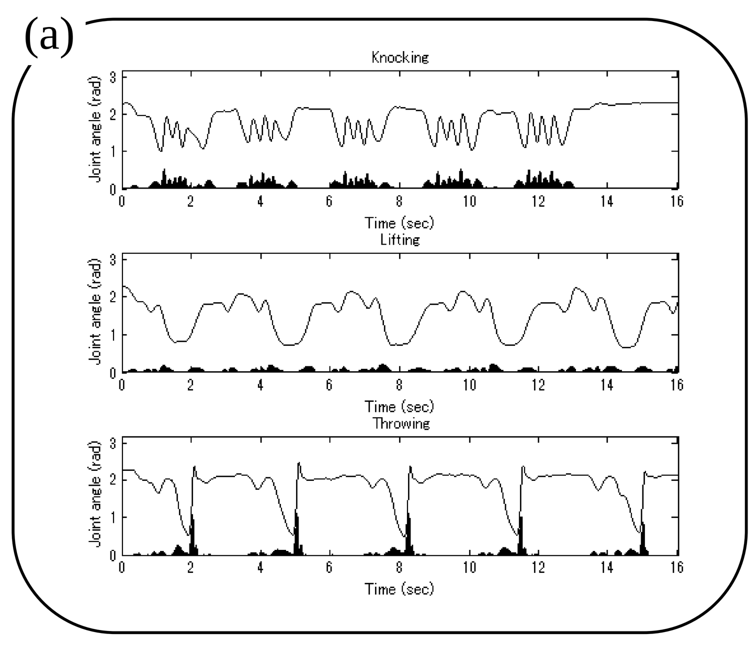


Fig 1: A hypothesized computational process of recognition of biological motion

Emotion recognition from action



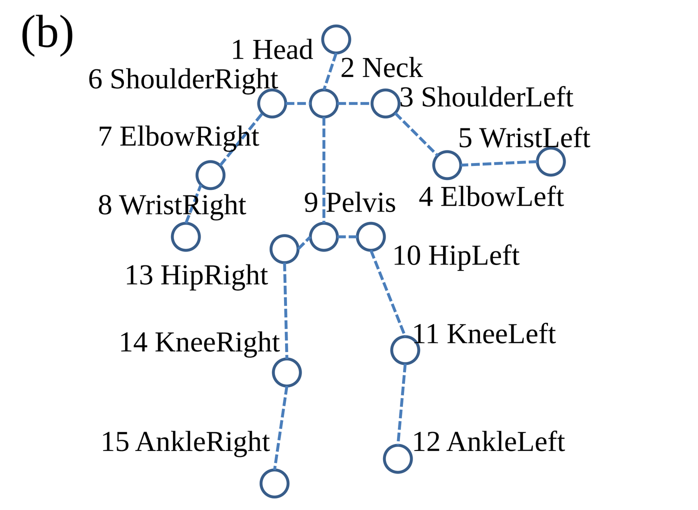
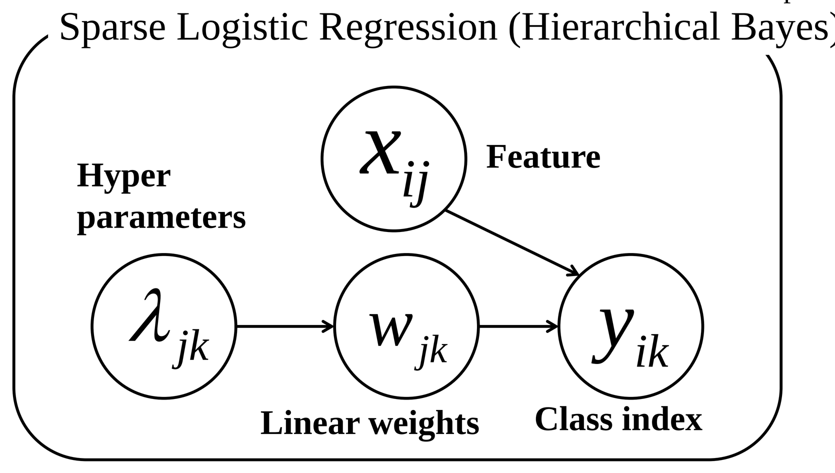


Fig 2: (a) Point-light actor (no link in the model and behavioral test), (a) A temporal profile of right-hand-elbow-shoulder joint angle (solid line) and its velocity profile (black) in 5 repeating knock, lift, and throw actions.

Sparse Logistic Regression (Hierarchical Bayes)



y_{ik} : **binary (0 or 1) class index of *k* (emotion/ action types) given trial *i***
x_{ij} : **information measure *j* given trial *i***
w_{jk} : **weights for the dimension *j* given class index *k* (normally varying).**
λ_{jk} : **hyper parameter (precision) of variability of weights.**

Fig 3: The sparse logistic regression linking the velocity features to given emotion/action class.

Behavioral study

- 10 subjects were asked to judge action and emotion types of presented point-light actors.
- *Stimuli*: Action-emotion stimuli were sampled from the biological motion library (Ma et al., 2006). Nine pairwise combinations of 3 actions (knock, lift, and throw) and 3 emotions (angry, happy, sad) were sampled from each of 3 selected actors. This yielded 27 video clips in total.

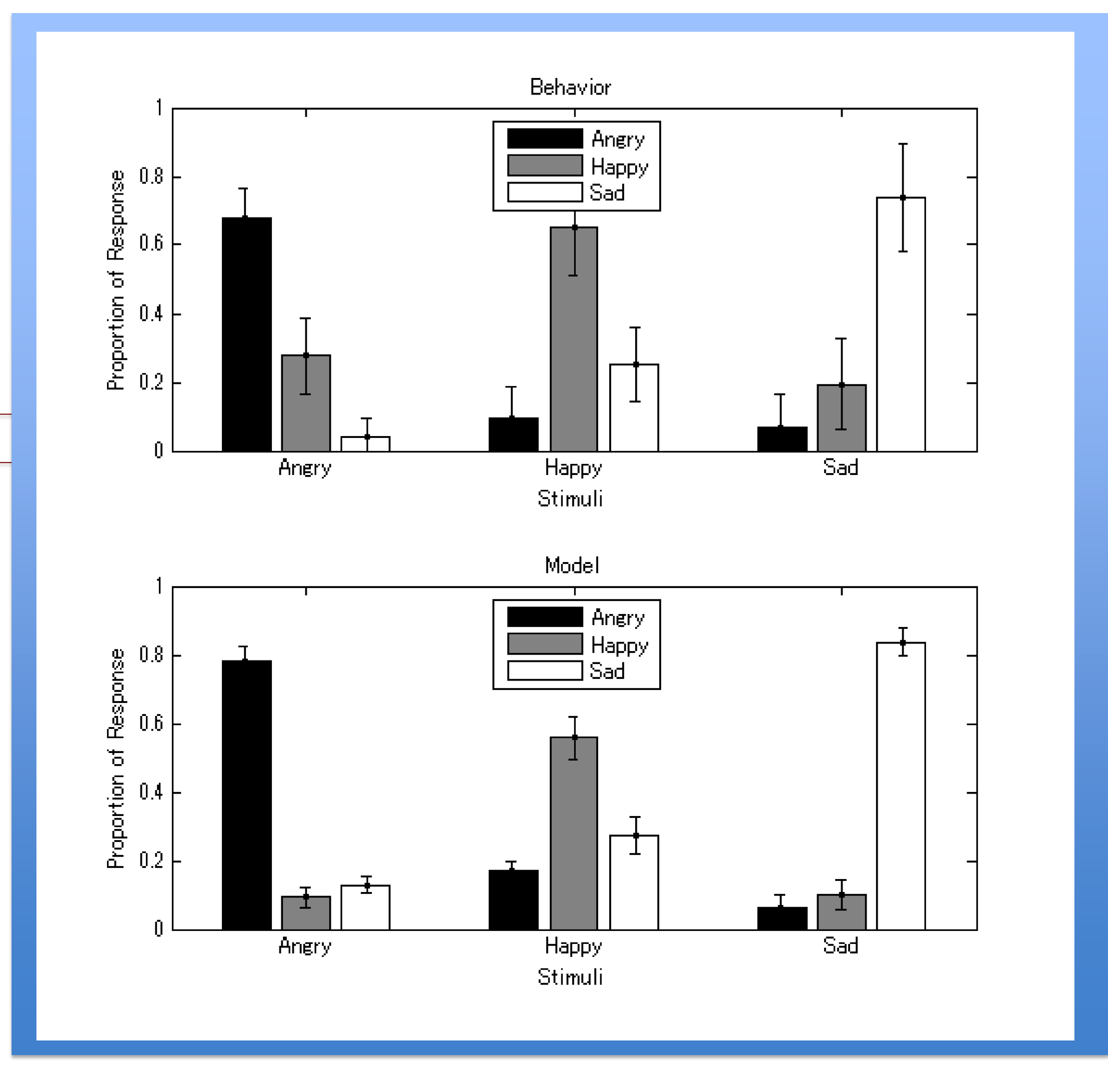


Fig 4. The response patterns for each emotion type in human subjects (upper panel) and the velocity model (bottom panel).

Results & Discussion

- Nested models with incrementally higher order features were compared *(0-DOF models)*.
- *Velocity* model: only velocity profile
- *Acceleration* model: velocity + acceleration profile,
- *Jerk* model: velocity + acceleration + jerk (third order derivative)

Acceleration or Jerk model fits better than Velocity model
Velocity (LL=-93.931/ R^2=0.810) < Acceleration (LL=-90.051/R^2=0.890)
< Jerk (LL=-89.116, -R^2=0.900) (but not A<J)

Action-specific recognizer does not improve fitting
Action-specific Acceleration model: -90.6381 (R^2=0.890)

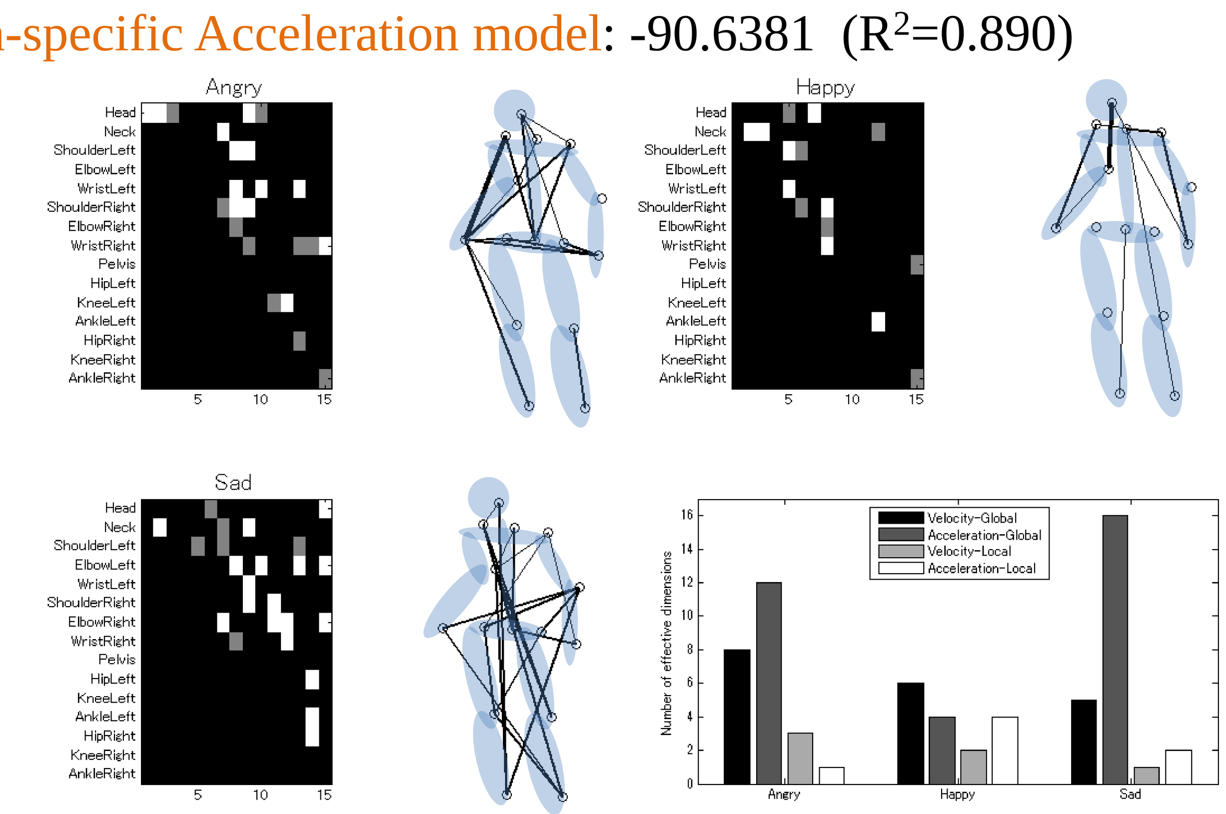


Fig 5. The variance/covariance in velocity profile significantly relevant to each emotion attribute mapped on a body scheme. The white and gray cell indicates effective variance/covariance of velocity and acceleration, respectively. No lower triangle cells were presented due to its symmetricity. The bottom right panel showed the number of effective dimensions for each emotion attribute.

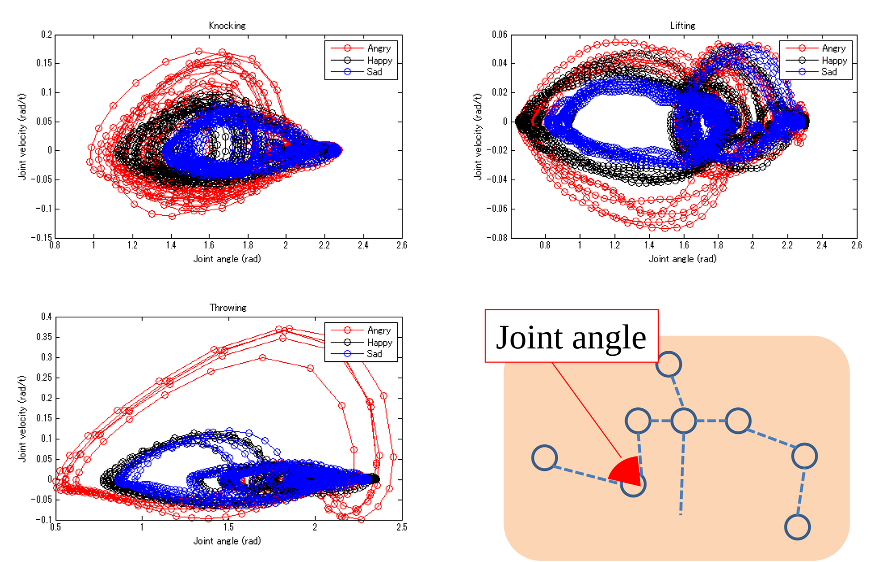


Fig 6. Joint space trajectories

Conclusions

Results: Emotional attributes in actions as well as action types could be identified by covariance of velocity profiles among multiple body parts.

Implication: Since, despite different velocity profiles in different actions, these features for emotional attributes were found commonly in multiple different actions, it suggests that the action styles may be mediated by an information channel parallel to action types per se.

References

Blake, R. & Shiffrar, M., (2007). Perception of Human Motion. Annual Review of Psychology, 58, 47–73.
Cook, J., Saygin, A. P., Swain, R., & Blakemore, S-H., (2009). Reduced sensitivity to minimum-jerk biological motion in autism spectrum conditions. Neuropsychologia, 47, 14, 3275-3278.
DeMeijer, M, (1989). The contribution of general features of body movement to the attribution of emotions. Journal of Nonverbal Behavior, 13, 4, 247-268.
Hobson, R. P. & Lee, A. (1999). Imitation and Identification in Autism, Journal of Child Psychological Psychiatry, 40, 4, 649-659.
Johhanson, G. (1973). Visual perception of biological motion and a model for its analysis. Perception & Psychophysics, 14, 2, 201-211.
Ma, Y., Paterson, H. M., Pollick, F. E. (2006). A motion capture library for the study of identity, gender, and emotion perception from biological motion, Behavior Research Methods, 38, 1, 134-141.
Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. Journal of Vision, 2, 371-387.
Troje, N. F. (2008). Biological motion perception. In Basbaum, A. et al. (Eds.), The senses: A comprehensive reference (pp. 231–238). Oxford: Elsevier.
Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: effects of structural and kinematic cues. Perception & Psychophysics, 67 (4), 667-675.
Yamashita, O., Sato, MA., Yoshioka, T., Tong F., Kamitani Y. (2008). Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns. Neuroimage. 42, 4, 1414-29.

This research was supported This study was supported by Artificial Intelligence Research Promotion Foundation and KAKENHI Grant-in-Aid for Scientific Research B No. 23300099.