Two interdisciplinary MSc projects in Machine Learning and Dance

Examiner: Hedvig Kjellström, hedvig@kth.se

Note that these projects are directed towards KTH students. We are not able to take in external students from other universities.

We are looking for two MSc students for an interdisciplinary collaboration with the Max Planck Institute for Empirical Aesthetics, Germany. The projects are each described in the following pages. Please contact Hedvig Kjellström if you are interested, indicating why you are applying, and specifying which project you would like to focus on.

In order to carry out these projects you need to have experience in Machine Learning and Computer Vision and a solid math foundation, e.g. corresponding to studies in the ML or SCR programs. If performed in a successful manner, the results will be publishable in an international peer-reviewed conference or journal.

Project 1: The Emotion in the Movement – categorizing the kinematics of emotions

Research question:
What are the kinematic features that predict different emotions in full body movements?

To address this question, we will use EMOKINE [1], a software and dataset creation framework for highly controlled kinematic datasets of emotionally expressive full-body movements. The primary novelty of this framework is that it provides the research community with a set of 12 readily computed kinematic features (32 statistics in total) that can be used out-of-the-box by researchers in order to model emotional expressivity in full-body movement. For the first time, these 12 are presented together: speed, acceleration, angular speed, angular acceleration, limb contraction, distance to centre of mass, quantity of motion, dimensionless jerk (integral), head angle (with regards to vertical axis and to back), and space (convex hull 2D and 3D).

A pilot dataset accompanies EMOKINE, recorded vith the XSENS® motion capture system. It consists of realistic full-body movement stimuli, which have been designed and instantiated with all other parameters controlled so that most of the variability stems from the different emotions expressed by the dancer. The data set comprises 9 sequences of 8 seconds length, danced 6 times, with 6 different emotional intentions by the dancer (joy, contentment, anger, fear, sadness, neutrality).

Task:
Firstly, we will investigate how well the emotion intended by the dancer can be recognized from the 12 features (32 degrees of freedom). Train a temporal classifier, e.g. based on LSTM [2], with 30-second snippets of kinematic features together with dancer emotional intent. Secondly, we will investigate how well the 12 features (32 statistic variables) correspond the features that a data-driven representation learning method train would learn to best classify the dancer emotional intent. Train a representation learning method, e.g. based on VAE [3], with a 32-dimensional latent space with the side constraint that the emotional intents should be as separated as possible. Compare this representation with the 12 EMOKINE features.

Download the dataset and all materials from Zenodo:

- https://zenodo.org/record/7821844

Download the EMOKINE software from Github:

- https://github.com/andres-fr/emokine

Comprehensive Readme files accompany the data and software.

References

[1] Christensen, J. F., Fernández, A., Smith, R. A., Michalareas, G., Yazdi, S. H. N., Farahi, F., . . . Roig, G. (2024). EMOKINE: A software package and computational framework for scaling up the creation of highly controlled emotional full-body movement datasets. *Behavior Research Methods*. doi:10.3758/s13428-024-02433-0

[2] Staudemeyer, R. C. (2019). *Understanding LSTM--a tutorial into long short-term memory recurrent neural networks*. arXiv preprint arXiv:1909.09586

[3] Yu, R. (2020). *A Tutorial on VAEs: From Bayes' Rule to Lossless Compression*. arXiv preprint arXiv:2006.10273

Project 2: Computational Dance Archeology – chasing the lost moves

Research question:
Can pose estimation models and temporal classification methods be used to reliably extract and classify dance poses from low-quality video materials of real dance performances from the 1960s and 70s?

After the political revolution of 1979, dance in Iran —both as everyday practice and as a cultural heritage— was first forbidden, and now remains heavily restricted. International, interdisciplinary research teams can contribute to safeguarding Iranian classical dance in the future. For example, by documenting the dance poses in archival video materials from before the revolution.

Our interdisciplinary team has produced a review of the dance style in collaboration with an Iranian dance specialist [1] and has created an Iranian dance movement library [2]. The movements and poses used in the video library are based on the book by Shahrzad Khorsandi (2015), who has systematically described the poses of the Iranian dance syllabus (see image below for examples).

Task:
We will work with a video dataset with annotated Iranian dance movements, and we also have access to more video data of similar dance styles, for model pretraining. Firstly, apply a SMPL [3] pose and shape model to extract the 3D pose in all videos (and evaluate the accuracy of the 3D estimation). Secondly, train an LSTM-based [4] classifier of these dance movements:

- Abshar & Qalammu → https://youtu.be/W5zn4yoru9c Time code: 4:41-4:58
- Segam foot pattern → https://youtu.be/w2Pr5v0wkY8 Time code: 4:59-5:08
- Shokufeh-Ahiz → https://youtu.be/AMzeG360_cg Time code: 0:57-1:13
- Additional examples: https://youtu.be/uT1LTY1ptQc

The movement library is available here:
- https://osf.io/2evgt/

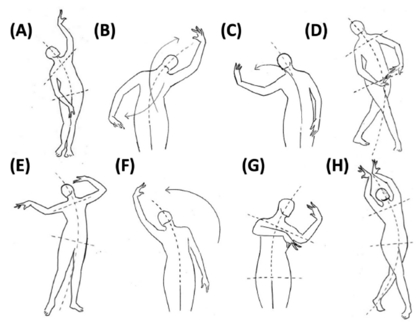

FIGURE 3 Examples of core movements in Iranian classical dance. Movements shown: (A) *Qalammu* (paintbrush), (B) *Ahiz* (pulling motion as with a sword being pulled from its scabbard), (C) *Raha* (release), (D) *Ranesh* (settling), (E) *Delroba* (heart-stealing), (F) *Shokufeh* (blossom), (G) *Nama* (showcase or view), and (H) *Parastu* (shorebird). Most of these movements can be used as connecting movements, or ending positions. The syllabus[10] contains 98 movements and positions. It has seven sections that define the categories of movement: foot, hand, and arm positions, hand movement, arm movement, isolations, foot patterns, turns, and figures (specific combinations of arm and leg movements and positions to present desired aesthetics and emotions). These movements are explored in various planes of motion and different bodylines, and analyzed in terms of the geometric relationship within the body. Much like Western classical ballet or Indian classical dance, which have a nomenclature and a movement vocabulary (e.g., *plié, tendu, passé retiré*), the movements of Iranian classical dance have been analyzed into their components. With the collaboration of an Iranian linguist, Dr. Kourosh Angali, as well as several other Iranians well versed in Farsi, a nomenclature was created (e.g., *Qalammu, Ahiz, Raha, Ranesh, Shokufeh, Delroba, Nama, Parastu*; the first five are arm motions and the last three are figures). Khorsandi is in a very unique position to offer such a syllabus. She has a master's degree in dance, has been focusing solely on Iranian dance since 1995, and was born and raised in Iran. She is today based in San Francisco, USA, and is a prominent figure in the Iranian diaspora dance movement. For the syllabus, Khorsandi assembled the core movements and positions of the dance style, for which no official systematization nor universal nomenclature had yet been published. It is important to note, however, that many dance practitioners have had their own systems of teaching and notation, without ever publishing them. We elaborate upon this issue in section "Systematization: Towards defining an elusive style." About her practice with this core movement language, Khorsandi says: "Sometimes, after performances, Iranians from the audience, especially older ones, approach me and say 'I've never seen Iranian dance performed this way before, but it really takes me back to Iran. Thank you so much!', which for me is an important marker that what I am doing is right."[154]

References

[1] Christensen, J. F., Khorsandi, S., & Wald-Fuhrmann, M. (2024). Iranian classical dance as a subject for empirical research: An elusive genre. *Annals of the New York Academy of Sciences*. doi:https://doi.org/10.1111/nyas.15098

[2] Christensen, J. F., Frieler, K., Vartanian, M., & Khorsandi, S. (2024, August 7). A Joy Bias: Emotion Perception from Full-Body Movements Is Modulated by Enculturation. *Retrieved from osf.io/2evgt*

[3] https://smpl.is.tue.mpg.de/

[4] Staudemeyer, R. C. (2019). *Understanding LSTM--a tutorial into long short-term memory recurrent neural networks*. arXiv preprint arXiv:1909.09586