DanceMoves: A Visual Analytics Tool for Dance Movement Analysis

Vasiliki Arpatzoglou*, Artemis Kardara*, Alexandra Diehl, Barbara Flueckiger, Sven Helmer, Renato Pajarola

University of Zürich, Switzerland

Abstract

Analyzing body movement as a means of expression is of interest in diverse areas, such as dance, sports, films, as well as anthropology or archaeology. In particular, in choreography, body movements are at the core of artistic expression. Dance moves are composed of spatial and temporal structures that are difficult to address without interactive visual data analysis tools. We present a visual analytics solution that allows the user to get an overview of, compare, and visually search dance move features in video archives. With the help of similarity measures, a user can compare dance moves and assess dance poses. We illustrate our approach through three use cases and an analysis of the performance of our similarity measures. The expert feedback and the experimental results show that 75% to 80% of dance moves can correctly be categorized. Domain experts recognize great potential in this standardized analysis. Comparative and motion analysis allows them to get detailed insights into temporal and spatial development of motion patterns and poses.

CCS Concepts

• ***Human-centered computing*** → ***Visual analytics;***

1. Introduction

The representation, storage, retrieval, and analysis of intangible assets, e.g. such as the choreography of dance, with the help of computers still poses several challenges [And16]. We agree with Sagasti's survey on the last fifty years in computer choreography [Sag19], in that there is still plenty of untapped potential when it comes to combining technology and art in this field. Calvert et al. put this down to dance being an art form that is rather slow in adopting technology, due to missing commercial opportunities and a certain reluctance of dancers and choreographers [CWRF05].

We focus on representing and visualizing dance moves to facilitate the study and analysis of dance poses and movement patterns. While there are textual descriptions used by choreographers, such as Labanotation [Gue05], these fall short when describing unorthodox moves in settings such as contemporary dance. In contrast, we envision a system that describes dance moves directly via the positions and movements of body parts, or groups of body parts. This is not entirely new, researchers have worked on how to capture (human) motion data before [WTLM04]. Nevertheless, several challenges remain, especially for analyzing motion data visually. Bernard et al. [BVKF17] identify three main challenges: First, the need for an underlying model that allows the effective cleaning and normalization of data, extraction of features, and definition of similarity measures. Second, the visualization and interaction design needs to support users in their task of analyzing and comparing motion data. Third, there should be a feedback loop with domain experts to improve the visualization tool and underlying model.

In this work, we address these challenges [BVKF17] and build a prototype of our framework *DanceMoves* for the interactive visual analysis of dance moves, with a focus on dance pose and motion quality analysis, as well as comparison and visual search of dance poses. In particular, the system offers the following functionality:

- Its visual and interactive overview depicts spatial, temporal, and velocity changes in body movements, allowing a user to detect and analyze repetitive poses globally, i.e, for the complete body, as well as locally, i.e., for specific areas or parts of the body.
- A similarity analysis to compare temporal and spatial changes in dance moves and poses qualitatively, i.e., visually, as well as quantitatively, i.e., with the help of different similarity metrics.
- A user can also formulate visual queries to search and locate similar dance poses in specific dance moves.

2. Related Work

Motion capture and analysis not only covers human motion but also applies to animals [WVZ*15] or vehicles [ST13], and other phenomena. To support domain experts in the area of dancing, we restrict ourselves to human motion patterns [DGL09, BWK*13, CGAG19]. Moreover, we are interested in interactive analysis of dance moves [ERSKI19, RTKI18]. Extending prior work in visual analysis of motion data [BWK*13, BVKF17, RESC16, WVZ*15] to the domain of dance and performing arts, we concentrate on three key aspects: direction and velocity of movement, as well as quality of poses. *ActionPlot* [CSS11] focuses on the analysis of

effort, tempo, intention, gaze, and balance through the use of a linear visualization of effort and tempo, enhanced by glyph encoding and color that describes the other variables. Although we also use glyphs and small multiples to show a temporal overview, we support visual comparison, visual search and quality analysis of dance poses. *Mova* [APS14] aims to visually analyze different human motion features such as speed and acceleration. The main difference again being that our tool allows for the visual comparison of dance moves, visual query and quality analysis of dance poses. Urribarri et al. present an overview+detail approach for comparison of Karate videos in [ULCP20]. Although we also use a heatmap grid in the overview visualization, our grid is interactive and acts as a dashboard to select one or more body parts, focus on one or more time frames, and perform partial and local analysis of dance moves.

3. Tasks Analysis

We performed an initial analysis of the requirements and tasks together with a ballet dancer. In a second iteration, we discussed our approach with two domain experts in film research and performing arts and theater. While gathering the main goals and tasks, we followed the tasks typology by Brehmer and Munzner [BM13]. The main goal is to provide immediate access to important features of dance moves, namely their direction and velocity patterns, while maintaining the context of the analysis at the frame level as well as the video level in one single overview. Based on this, we defined three main visual tasks: providing an overview of the dance moves, visually comparing videos, and searching for specific dance poses.

Task 1: *Overview and Analysis of Dance Moves Patterns.* The first task is usually to get an overview of the dance moves in a video. First, the user can analyze direction and velocity changes on a *video level* and then do so for a specific dance pose at the *frame level*. This can be done *globally*, for the whole body, or *locally*, for certain body parts. Based on this, dancers, choreographers or dance and film scholars can assess and analyze poses and motion patterns at different levels of detail, such as locating the continuous extension and retraction of legs and arms.

Task 2: *Comparison of Dance Moves and Assessment of their Quality.* In this task, a user compares two dance movements over time. This analysis allows the user to understand the similarities and dissimilarities between two dance moves, and it can be done at local or global scale, and at frame level or video level. In this way, the user can determine the quality of the movements and how much they differ when performed at various speeds.

Task 3: *Search for Specific Dance Poses.* In this task, a user has a specific dance pose in mind and wants to search for it at the video level to analyze its occurrence and periodicity. This analysis can be done globally for the whole body, or locally for specific body parts, and will locate the dance pose in a video.

4. DanceMoves Solution

Fig. 1 illustrates the main parts of our solution: the frontend (see Sec. 4.1), composed of the visual components, and the backend (see Sec. 4.2), composed of an estimation model and similarity metrics.

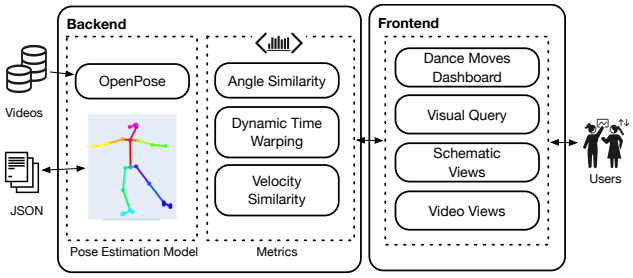

Figure 1: *Architecture with backend and frontend components.*

4.1. Visual Design

For the visual design of our solution, we selected a Multiple Coordinated Views layout [Rob07] that allows the user to always have a complete overview of the main visualization components and interact with them. We designed the visual components to support Tasks 1, 2, and 3, as described in Section 3.

4.1.1. Overview and Analysis of Dance Moves

We designed a *dance moves dashboard* to provide an interactive overview and easy manipulation of dance poses, motion changes over time, and body areas. The first column in Fig. 2(c) shows small multiples of different body parts, the first row shows video frames, and the internal cells show different values depending on the configuration. Additionally the video is shown, see Fig. 2(a), with a slider to enable the user to interactively select a given frame and highlight it in the dance moves dashboard. The detected segments, linked to the video, are shown in a schematic view in Fig. 2(b).

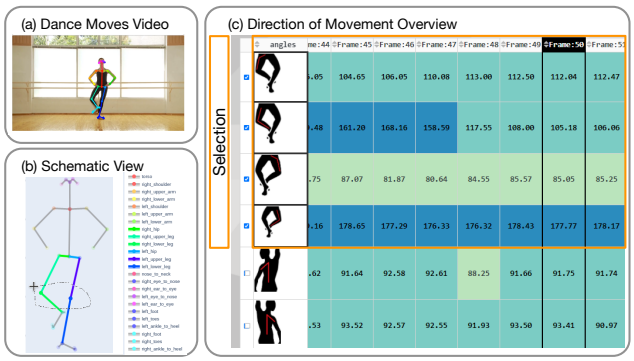

Figure 2: *Dashboard overview: (a) video section for interaction with the video, (b) schematic view corresponding to the selected frame, and (c) dance moves overview for selecting one or more body parts and observing movement changes over time.*

4.1.2. Visual Comparison and Similarity Analysis

For the visual comparison of two dance poses, the dashboard is adapted to present an overview based on the *Motionrugs* concept [BJC*18], allowing for detection of large changes. Each video can have a different length, so we align the dance poses using DTW (explained in Sec. 4.2), and visually highlight the aligned dance

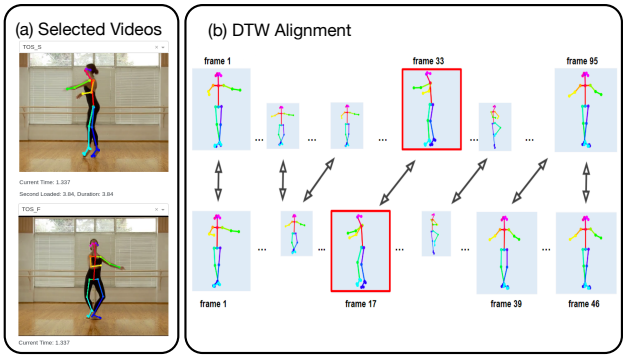

Figure 3: *DTW alignment according to the similarity metric.*

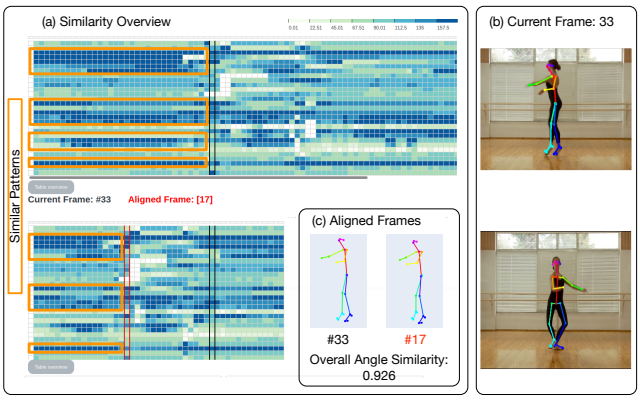

Figure 4: *Visual comparison: (a) Selection of two videos, (b) color coded similarity values, and (c) schematic views. (d) pairs of videos aligned according to DTW algorithm.*

poses in red as shown in Fig. 4(a), corresponding to the video frames in Fig. 4(b). The two schematic views in Fig. 4(c) allow a user to maintain the context of the current dance poses, and visually assess how they differ topologically, at different body parts.

4.1.3. Visual Search and Dance Pose Quality Analysis

The mini-thumbs schematic view in Fig. 5(c) supports quick visual search for common dance poses in the video. The *dance moves dashboard* in Fig. 5(d) is used as a pivot table where the user can look for complete or partial dance pose matches. The second dashboard column always shows the directions of movement corresponding to the queried dance pose. The rest of the dashboard shows the dance pose dissimilarities, frame by frame. The same exact pose as selected in the video is highlighted as a complete match. When it is partially detected, i.e., only some parts of the body are similar to the selected pose, it is highlighted as a partial match.

4.2. Similarity Measures

At the backend, we started by defining a similarity measure for the static case, i.e., comparing individual frames of different videos to each other and computing their similarity. First, we extract features from these frames. We utilize the Body-25 pose estimation model that extracts the (two-dimensional) positions of 25 parts of the body, such as left shoulder, right knee, or neck. This model is included in the official OpenPose real-time multi-person keypoint detection system [SJMS17] [CHS*19] [CSWS17] [WRKS16].

After extracting the features from a frame, we make sure that they are comparable by normalizing them. We did not use the positions of the body parts directly but defined segments using the positions as endpoints of these segments. The stick figure under the box 'OpenPose' in Fig. 1 illustrates this, i.e., the yellow segment for the lower right arm has as endpoints the positions of the right hand and the right elbow. Moreover, we do not use the segments directly but measure the angles between adjacent segments, giving us a vector of 29 angles describing a pose. This description is independent of the relative size and position of the dancer within the frame. Given two vectors of angles, we determine the distance between two dance poses via the cosine measure. We call this the *angle similarity*, and it is used as a measure of relative direction of the movement. We can compute the *angle distance* by subtracting the angle similarity (which is a value between 0 and 1) from 1. According to the domain experts, the velocity and fluidity of dance moves can also play a role when comparing two videos (especially for contemporary dance). Thus, we define a second similarity measure to capture the velocity of the movement. We use the differences between angles from one vector to the next. For example, if $v_1 = (20, 10, \ldots, 70)$ and $v_2 = (15, 20, \ldots, 50)$, then $\Delta = v_2 - v_1 = (-5, 10, \ldots, -20)$. Clearly, this only works when we have at least two frames, and a video sequence consisting of n frames is described by $n-1$ difference vectors. We call this the *velocity similarity*. Like before, the *velocity distance* can be computed by subtracting the velocity similarity from 1.

In a second step, we generalize the similarity measure to the dynamic case, i.e., computing the similarity between two video sequences consisting of multiple frames. As the moves of two dancers are not perfectly synchronized even when they perform the same sequence of dance poses, we need to align the frames in the two sequences to find the best match. We utilize dynamic time warping (DTW) to align the frames, which minimizes an alignment cost. DTW has proven valuable for aligning non-linearly and irregularly shifted time series data [SSS*20], including human motion [CCW*04, RTKI18, ULCP20]. A limitation of DTW is that it does not normalize the costs. The costs of the individual steps are just added up and the total value can vary widely depending on the lengths of the sequences. Therefore, we use DTW to determine the least costly alignment of the frames and then compute a normalized distance (or similarity) between two videos by determining the weighted average distance (or similarity) between all aligned pairs on the minimal-cost warping path (see Fig. 3).

5. Illustrative Use Cases and Domain Experts' Feedback

Our tool was evaluated using the "MultiTime Laboratory H-Dance Database" [VSG*14, SV14], consisting of 83 short videos containing different ballet poses with different movement variations, but using only 73, as not all would play properly.

5.1. Understanding Dance Movement Structures

For domain experts, structures of movements are indicative of the motion patterns, their quality and assessment. We illustrate the use of our tool by analyzing a *frappé* dance movement in Fig. 2, where the legs are the most important body parts, as it consists of an on-the-spot extension of the leg to the front followed by a kick of the leg to the side and back. A user can select the body parts that are of interested and analyze them by the *dance moves dashboard*. The selected body parts are highlighted in all coordinated views. While a user may be able to follow simpler movements by just watching the video, our tool allows them to keep track of many different groups of body parts simultaneously, providing the relevant information at a glance, see Sec. 4.1.1 for details.

5.2. Similarity Analysis over Time

Comparing two movements over time is not an easy task, as they may be executed at different speeds and shifted within the videos. However, similarity metrics are very important for advanced search and comparison functions applied to dance videos. For example, Fig. 4 shows two videos where the dancer is performing the same dance movement, namely a *tour piqué*. The frames in Fig. 4(b) are currently temporally aligned, i.e., they both show frame 33, which due to the different duration of the videos show different poses. The dashboards depict two pixel tables with angle values. Similar shades indicate a constant pattern of movement and varying shades show changed patterns, associated with motion. Exception is the white color which indicate undetected angles or invisible ones. A time-warped alignment is applied, as described in Sec. 4.2. Fig. 3(a) shows the selected frame in each compared video. Fig. 3(b) shows how with the help of DTW, the frames can be aligned according to their similarity. Fig. 4(c) shows the matching poses at frames 33 and 17 respectively.

5.3. Searching a Pose in Video / Visual Query

Locating dance movements and specific poses in videos give domain experts immediate access for the investigation of specific dance poses and their execution. Fig. 5(c) shows three different poses (*first arabesque*, *demi-plié*, *relevé*) and we assume that the user chooses the *relevé* dance pose. The first column in Fig. 5(d) contains all the angle values of all body parts we would expect for a dancer performing a *relevé*, all the other columns show the differences to these values for every frame in the video. When the difference between them is zero, the color is white. As shown at the sixth frame, the differences for all body parts are zero and this is a complete match. A partial match later on shows that the dancer repeats the *relevé* pose only for certain body parts, namely head, hands, arms and torso. Our tool allows a user to instantly locate complete and partial matches of dance poses in videos, highlighting all the relevant information, as explained in Sec. 4.1.3.

5.4. Experts' Feedback

Domain experts found our approach useful to analyze reoccurring poses over time. The similarity analysis allows to make connections and similarities among body parts movements and positions

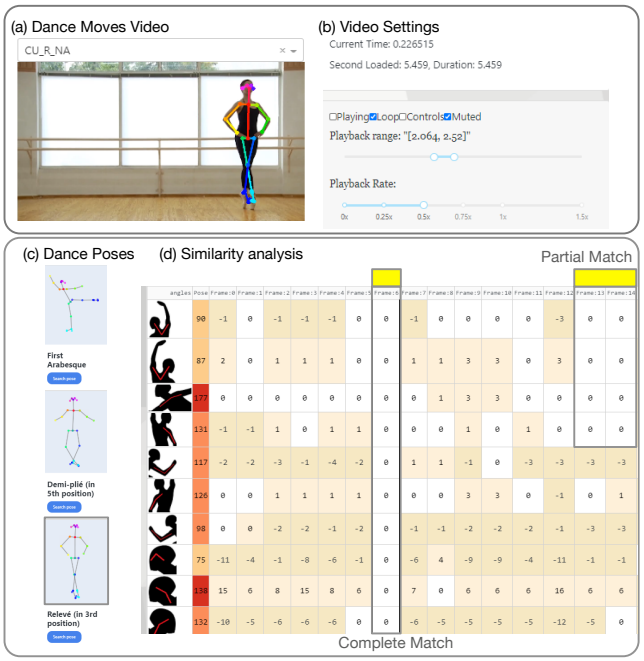

Figure 5: *Dance Pose Analysis: (a) the video, (b) its settings, (c) the selected dance pose, and (d) vector describing the dance pose (1st column) and dissimilarities per frame (other columns).*

for the immediate analysis and assessment of dance moves. They also highlighted that our tool helps them to investigate specific patterns that may occur between and after the pose has been performed. Current features could be extended to automatically detect repetitive poses in a video. The tool enables them to gain insights into poses, motion patterns and development over time both on the level of single poses, on the level of the whole dance videos and even by the comparison of different videos.

6. Conclusions and Future Work

DanceMoves is a visual analytics framework for analysis of dance moves, direction of movement and velocity, similarity measures, and visual query and quality assessment of dance poses. To evaluate the similarity measures, we applied hierarchical agglomerative clustering to the videos in our test data. With angle similarity, 19 out of 73 videos ended up in the wrong cluster, whereas with velocity similarity, 15 of the videos were misclustered. These are promising early results, with a large majority of correctly clustered videos. Regarding the visual design, domain experts found the tool useful to gain insights into individual performances, and to establish and evaluate hypotheses through the visualizations and similarity metrics. More integration of screenshots, motion trajectories, integration of sound analysis and intuitive interactions with the results would be an asset in future developments of the tool set.

7. Acknowledgements

We thank Dr. Gunter Lösel for his valuable feedback to our project.

References

[And16] ANDERSON D. P.: Nailing smoke – curation at the bleeding edge of technology. *Communciations of the ACM 59*, 12 (December 2016), 37–39. doi:10.1145/3012423. 1

[APS14] ALEMI O., PASQUIER P., SHAW C.: Mova: Interactive movement analytics platform. In *Proceedings International Workshop on Movement and Computing* (2014), pp. 37–42. doi:10.1145/2617995.2618002. 2

[BJC*18] BUCHMÜLLER J., JÄCKLE D., CAKMAK E., BRANDES U., KEIM D. A.: Motionrugs: Visualizing collective trends in space and time. *IEEE Transactions on Visualization and Computer Graphics 25*, 1 (January 2018), 76–86. doi:10.1109/TVCG.2018.2865049. 2

[BM13] BREHMER M., MUNZNER T.: A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (October 2013), 2376–2385. doi:10.1109/TVCG.2013.124. 2

[BVKF17] BERNARD J., VÖGELE A., KLEIN R., FELLNER D. W.: Approaches and challenges in the visual-interactive comparison of human motion data. In *Proceedings International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* (2017), pp. 217–224. doi:10.5220/0006127502170224. 1

[BWK*13] BERNARD J., WILHELM N., KRÜGER B., MAY T., SCHRECK T., KOHLHAMMER J.: Motionexplorer: exploratory search in human motion capture data based on hierarchical aggregation. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (December 2013), 2257–2266. doi:10.1109/TVCG.2013.178. 1

[CCW*04] CHIU C.-Y., CHAO S.-P., WU M.-Y., YANG S.-N., LIN H.-C.: Content-based retrieval for human motion data. *Journal of Visual Communication and Image Representation 15*, 3 (September 2004), 446–466. doi:10.1016/j.jvcir.2004.04.004. 3

[CGAG19] CASERMAN P., GARCIA-AGUNDEZ A., GÖBEL S.: A survey of full-body motion reconstruction in immersive virtual reality applications. *IEEE Transactions on Visualization and Computer Graphics 26*, 10 (December 2019), 3089–3108. doi:10.1109/TVCG.2019.2912607. 1

[CHS*19] CAO Z., HIDALGO MARTINEZ G., SIMON T., WEI S., SHEIKH Y. A.: Openpose: Realtime multi-person 2D pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (May 2019). doi:10.1109/TPAMI.2019.2929257. 3

[CSS11] CARLSON K., SCHIPHORST T., SHAW C.: Actionplot: A visualization tool for contemporary dance analysis. In *Proceedings of the International Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging* (2011), pp. 113–120. doi:10.1145/2030441.2030466. 1

[CSWS17] CAO Z., SIMON T., WEI S.-E., SHEIKH Y.: Realtime multi-person 2D pose estimation using part affinity fields. In *Proceedings Computer Vision and Pattern Recognition* (2017). doi:10.1109/CVPR.2017.143. 3

[CWRF05] CALVERT T., WILKE W., RYMAN R., FOX I.: Applications of computers to dance. *IEEE Computer Graphics and Applications 25*, 2 (March 2005), 6–12. doi:10.1109/MCG.2005.33. 1

[DGL09] DENG Z., GU Q., LI Q.: Perceptually consistent example-based human motion retrieval. In *Proceedings Symposium on Interactive 3D Graphics and Games* (2009), pp. 191—-198. doi:10.1145/1507149.1507181. 1

[ERSKI19] EL RAHEB K., STERGIOU M., KATIFORI A., IOANNIDIS Y. E.: Dance Interactive Learning Systems: A Study on Interaction Workflow and Teaching Approaches. *ACM Computing Surveys 52*, 3 (June 2019), 1–37. doi:10.1145/3323335. 1

[Gue05] GUEST A. H.: *Labanotation – The System of Analyzing and Recording Movement*. 2005. doi:10.4324/9780203626122. 1

[RESC16] RAGAN E. D., ENDERT A., SANYAL J., CHEN J. J.: Characterizing provenance in visualization and data analysis: An organizational framework of provenance types and purposes. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (August 2016), 31–40. doi:10.1109/TVCG.2015.2467551. 1

[Rob07] ROBERTS J. C.: State of the art: Coordinated & multiple views in exploratory visualization. In *International Conference on Coordinated and Multiple Views in Exploratory Visualization* (2007), IEEE, pp. 61–71. doi:10.1109/CMV.2007.20. 2

[RTKI18] RAHEB K. E., TSAMPOUNARIS G., KATIFORI A., IOANNIDIS Y.: Choreomorphy: A whole-body interaction experience for dance improvisation and visual experimentation. In *Proceedings International Conference on Advanced Visual Interfaces* (2018), pp. 1–9. doi:10.1145/3206505.3206507. 1, 3

[Sag19] SAGASTI F.: Information technology and the arts: The evolution of computer choreography during the last half century. *Dance Chronicle 42*, 1 (April 2019), 1–52. doi:10.1080/01472526.2019.1575661. 1

[SJMS17] SIMON T., JOO H., MATTHEWS I., SHEIKH Y.: Hand keypoint detection in single images using multiview bootstrapping. In *Proceedings Computer Vision and Pattern Recognition* (2017). URL: http://arxiv.org/abs/1704.07809. 3

[SSS*20] SOLTESZOVA V., SMIT N. N., STOPPEL S., GRÜNER R., BRUCKNER S.: Memento: Localized time-warping for spatio-temporal selection. In *Computer Graphics Forum* (2020), vol. 39, pp. 231–243. doi:10.1111/cgf.13763. 3

[ST13] SIVARAMAN S., TRIVEDI M. M.: Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis. *IEEE Transactions on Intelligent Transportation Systems 14*, 4 (July 2013), 1773–1795. doi:10.1109/TITS.2013.2266661. 1

[SV14] SGOURAMANI H., VATAKIS A.: "Flash" dance: How speed modulates perceived duration in dancers and non-dancers. *Acta Psychologica 147* (March 2014), 17–24. doi:10.1016/j.actpsy.2013.06.009. 3

[ULCP20] URRIBARRI D., LARREA M., CASTRO S., PUPPO E.: Visualization to compare karate motion captures. In *Congreso Argentino de Ciencias de la Computación* (2020), pp. 446–455. doi:10.1007/978-3-030-48325-8_10. 2, 3

[VSG*14] VATAKIS A., SGOURAMANI H., GOREA A., HATZITAKI V., POLLICK F. E.: Time to act: New perspectives on embodiment and timing. *Procedia - Social and Behavioral Sciences 126* (March 2014), 16–20. doi:10.1016/j.sbspro.2014.02.302. 3

[WRKS16] WEI S.-E., RAMAKRISHNA V., KANADE T., SHEIKH Y.: Convolutional pose machines. In *Proceedings Computer Vision and Pattern Recognition* (2016), pp. 4724–4732. doi:10.1109/CVPR.2016.511. 3

[WTLM04] WEIMING HU, TIENIU TAN, LIANG WANG, MAYBANK S.: A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on Systems, Man, and Cybernetics, Part C 34*, 3 (July 2004), 334–352. doi:10.1109/TSMCC.2004.829274. 1

[WVZ*15] WILHELM N., VÖGELE A., ZSOLDOS R., LICKA T., KRÜGER B., BERNARD J.: FuryExplorer: visual-interactive exploration of horse motion capture data. In *Proceedings Visualization and Data Analysis* (2015), vol. 9397, pp. 148–162. doi:10.1117/12.2080001. 1