Motion, Captured: an Open Repository for Comparative Movement Studies

Varsha Iyengar
School of Arts, Media & Engineering
Ira A. Fulton Schools of Engineering
Arizona State University
Tempe, AZ 85281, USA
varsha.iyengar@asu.edu

Grisha Coleman
School of Arts, Media & Engineering
Arizona State University
Tempe, AZ 85281, USA
grisha.coleman@asu.edu

David Tinapple
School of Arts, Media & Engineering
Arizona State University
Tempe, AZ 85281, USA
david.tinapple@asu.edu

Pavan Turaga
School of Arts, Media & Engineering
School of Electrical Computer & Energy Engineering
Arizona State University
Tempe, AZ 85281, USA
pavan.turaga@asu.edu

ABSTRACT

This paper begins to describe a new kind of database, one that explores a diverse range of movement in the field of dance through capture of different bodies and different backgrounds - or what we are terming movement *vernaculars*. We re-purpose Ivan Illich's concept of 'vernacular work' [11] here to refer to those everyday forms of dance and organized movement that are informal, refractory (resistant to formal analysis), yet are socially reproduced and derived from a commons. The project investigates the notion of vernaculars in movement that is intentional and aesthetic through the development of a computational approach that highlights both similarities and differences, thereby revealing the specificities of each individual mover. This paper presents an example of how this movement database is used as a research tool, and how the fruits of that research can be added back to the database, thus adding a novel layer of annotation and further enriching the collection. Future researchers can then benefit from this layer, further refining and building upon these techniques.

The creation of a robust, open source, movement lexicon repository will allow for observation, speculation, and contextualization - along with the provision of clean and complex data sets for new forms of creative expression.

MOCO'16, July 5–6, 2016, Thessaloniki, Greece.

Author Keywords

Movement; motion capture; annotation; repository; database; open-source; crowd-sourcing; translation; community; machine-learning; clustering.

ACM Classification Keywords

H.1.2 Information Systems: Models and Principles – User/Machine Systems: H.5.2

Information interfaces and presention:Multimedia Information Systems–User Interfaces

INTRODUCTION

In his 1980's book *Vernacular Values*, Ivan Illich defines vernacular work as work done outside any recognized market, [...] unpaid activities which provide and improve livelihood, but which are totally refractory to any analysis utilizing concepts developed in formal economics. He argues that even in late industrial societies, a large majority of all goods and services are vernacular in nature, functioning in a shadow economy where money is not the currency. We have extended this conceptual framework to the time-based movement arts that, as an 'object', in contrast to other art forms, has heretofore been resistant to methods of commodification, remaining in the realm of specialized forms of written movement notation [9]. While still image, video, and audio recording technologies have accelerated the means and ways in which those media are produced and innovated – movement as digital object is a relatively new phenomenon. Current powerful motion capture technology enables the capture and representation of complex human movement with a density of details in ways that have been impossible until recently [22].

As humans, our capacity to recognize and distinguish amongst different kinds of movement (animal, human, animate, inanimate) is a foundational evolutionary ability [14], as basic as breathing, and rarely reflected upon unless that ability is damaged or lost. The patterns of movement in an individual mover can often be immediately recognized from extraordinarily small and sparse sets of data. Leveraging the affordances of both the human ability to recognize and classify movement, along with the rich data that the MoCap system provides, this project explores how movement identities can be an emergent property of data analysis.

However, given the context for the development of such advanced technologies for movement capture, an anthropological critique of the *kinds of movement* to be captured is a part of this project's thematic goals.Motion capture data is now available as open and un-annotated [5]. Yet, we observe an underlying "computational perspective" in which the digital ~~representation of the~~ skeletal ~~form~~ is represented as a relatively *neutral* picture of human movement - to be applied across multiple uses. This project ~~challenges this notion~~ of digital neutrality. In the context of dance studies, Performance Studies scholar Susan Manning ~~has raised the point~~ that; "[...] the history of anthropology has raised the larger question of how ethnographers from "First World" countries came to know and interpret the cultures of "Third world" countries [...]" [17]. ~~Western histories of science and technology have evolved in tandem with European classical forms of dance - which have been culturally connected to concepts and values of lines and planes in Cartesian space~~ [6], ~~contributing to the notation systems that ensued.~~ Culturally complex movement forms that have emerged from vernacular styles such as Hip-hop and Bharathanatyam and other diasporic movement traditions have not typically been represented through written notation systems ~~[e.g. objects]~~ for a host of cultural and technical reasons. This project asks; *what might be revealed in recording, mining and comparing distinct movement vernaculars? What could a platform for the development of a digital movement lexicon that considers context along with capture?*

This paper investigates one approach to the analysis of patterns in human movement, in the sense of being discoverable by standard clustering algorithms (k-means, sequential k-means) through data. We use these clustering tools to detect repeated patterns in movement; the postures that tend to reappear, often unintentionally, as markers of an identity or 'movement fingerprint'. This research focuses on understanding the statistical distributions that constitute an individual's motion data, magnifying the commonalities and differences that constitute repeated [habitual] patterns, thereby exposing the unique vernacular of the mover. Notably, this project puts particular emphasis on the process of curating the database through choosing source movement from a stylistic and culturally diverse community of movers, before the technological capture process begins.

While other databases of human movement certainly exist (gesture, pedestrian, dance), technological procedures in current data motion capture are, by definition, processes of reduction and separation. This project explores approaches to sourcing and capturing movement data that incorporates and foregrounds the environmental, cultural, technological, economic, and historical contexts in which it exists. As a repository for a growing collections of high-resolution motion capture 'portraits' – the goal of the repository is in becoming a community resource that allows stakeholders (researchers, theorists, creatives) to annotate and add layers of new content. The portraits themselves are comprised of synchronized sound, video, and 3D motion-capture data with a wide variety of styles and from a range and diversity of cultural contexts. This project aims to provide a simple and accessible platform for creative interpretation, translation and annotation of rich, complex, and non-'noisy' movement data. The repository can then act as a public space for a range of artists, researchers, dancers, ethnographers, humanists, and somatic movement educators to respond and add novel layers of object creation, social and historical context, and technological and somatic analysis. In building an expandable platform that can embed multiple disciplinary perspectives, this project considers the development of a movement lexicon that includes form and context through collection, analysis, and comparison of distinct movement vernaculars. This paper describes a single research path as an example of how we have begun to build and use the repository.

move the green bit down

RELATED WORK

Motion Capture and Movement Repositories

There currently are several examples of open source libraries of motion capture data [5]. These tend to focus on either simple useful gestures (running, walking, jumping) intended for use by game developers, or on motion capture of complex choreographed forms of professional dance. Motion capture data sets such as the *Open Motion Project* at ACCAD [2] set out to make available "... a set of motions that are useable for motion for video games, animation, etc." Working on the other side of the spectrum, *Motionbank.org* [19] has created a "... network of choreographers, dancers and researchers interested in using MoSys for their own purposes", with the stated aim of "translating choreography and dance into new digital forms."

A notably different project leveraging motion data is a project by Evan Roth called *White Glove Tracking* [18], in which he creates an experimental framework for not only annotating the motion in a video, but then also providing a community platform to highlight community responses. While this work is in itself an art project, it serves as an example of the wealth of novel and creative responses that can emerge from a community. Structurally similar, the well known collaboration between William Forsythe, ACCAD and OSU Dance, *Synchronous Objects* makes available not only the score but multi-angle video data of a performance of *One Flat Thing, Reproduced* by William Forsythe, to a community of visualization experts and artists to re-interpret. The results of this community collaboration include stunning examples of video abstraction, statistical analysis, data visualization, architectural forms, generative drawing, and 3D animation all derived from the original score and performance data. Synchronous Objects and the White Glove Tracking Project both point towards the incredible latent potential in forming multi-modal representations of data as well as, in terms of the White Glove

Tracking Project, crowd sourcing data sets related to human movement and dance. Another repository that provides storage, visual browsing and annotation of motion capture data set is the RepoVizz [1]. It has a highly advanced user interface that allows easy navigation through a tree-based structure of multi-modal data streams and rendering of real-time data through WebGL. In this project we aim to direct the power of community analysis not at a single highly choreographed performance but rather at the wide range of vernacular movements that we encounter around the world in our everyday lives.

Computational Systems for Human Movement Analysis

The use of statistical machine learning in motion analysis forms a significant field of interest in the movement and computing community. Françoise et al. [13] proposed using Hidden Markov Regression as a methodology for movement sequence analysis to investigate performers' consistency with a use case in Tai Chi. Barbic et al. [3] proposed an algorithm to decompose human motion and places a cut when the distribution of human poses changes based on probabilistic principal component analysis. We are not interested in recognition or decomposition of gesture itself, but in visualizing dominant clusters. We put forth this visualization of clusters as an example of a new layer of annotation, added back into the database. A variety of clustering approaches have been applied to movement data and video [21][16]. Zhou et al., [25] uses Hierarchical Aligned Cluster Analysis to temporally cluster human motion in motion capture data and video. In the interest of computational efficiency, we choose to use k-means and an online variant of k-means clustering, which allows processing of movement data as it arrives, and updates clusters on the fly. This approach to analysis also allows for use in interactive or live performance contexts. One approach to this interaction is having the system learn from the dancer, thus giving the dancer directional control of the computational system [10]. A different approach focuses on the complex collaborative performer-machine relationship to build a deeper conversation between the mover and computational system. John McCormick's *Emergence* [12] uses Artificial Neural Networks to create a digital agent that learns to dance through a dancer creating a semi-improvised interactive visual performance. Such analysis approaches can add to creating a stronger machine/performer relationship that can be used for such real-time decomposition of dance.

SYSTEM HARDWARE

We designed our initial motion capture sessions to capture two streams of data in parallel; a primary stream consisting of optical motion capture using a commercial IR OptiTrack motion capture (MoCap) system and HD video recordings. The OptiTrack uses a 12-camera system running the "Motive: Body" real-time analysis/recording software. The mover wears a suit with 37 retro-reflective markers and is analyzed to determine the full-body skeleton in real-time at 120 fps. For our capture design, the dancer performs several short movement phrases, and MoCap data is exported in a variety of formats (C3D/BVH/CSV). The additional three HD cameras record video of the subject from three key angles – left, center and above – while also recording audio of the session. The video is trimmed and uploaded to a server to begin analysis.

METHODS OF DATA COLLECTION

Movement data was captured from practitioners of different dance styles including contemporary African-American and European vernaculars, Hip-hop and Bharathanatyam. Simple prompts with parameters around axes of rhythm, space, and duration were created, keeping in mind differences in approach and execution of styles. In designing a set of these prompts or instructions for the dancer, we were sensitive to creating prompts that were simple and open enough encourage improvisation, while allowing us to create a framework for structured similarity across movers. This 'open structure' then allows us to find a common basis for further analysis and organized the capture process. The motion data was then processed to remove individual marker occlusions, resulting in very clean data representing major joints of the skeleton of the dancer. The final representation is a point-cloud of 3D locations of the joints, with the hip joint considered as the origin.

COMPUTATIONAL PATTERN DISCOVERY

Generally speaking, somatic education emphasizes creating conditions for more efficient, functional movement patterns to emerge [8][24]. Yet it is also understood that patterns of habitual movement are, at any given moment in a person's life, an accumulation of who where and what the person has done and where they have been in their lives [4]. Our research shows that an individual's repetitive patterns of movement are often extraordinarily 'high fidelity' – in that the digital representation of that movement can be successfully recognized with small amounts of data. Analysis methods using clustering techniques can extract movement patterns from a sample set of dancers to reveal distinct 'signatures', and further, explore what this type of analysis tells about their particular movement. We do this with the clustering techniques of k-means, along with a real-time clustering technique, sequential k-means.

Clustering via k-means

Consider the recording of movement sequences represented as vectors $(x_1, x_2, \ldots, x_n)$, where x_i is the i^{th} vectorized posture. K-means clustering is an unsupervised learning approach that, informally, aims to partition the n observations into k groups by minimizing the *L2* error between a set of centroids and the data points associated to each centroid. [7]

Our reasons for using k-means as the clustering technique is two-fold. It opens the possibility for extending the approach to a real-time setting via variants such as sequential k-means. We elaborate on this method for a work-in-progress performance later in this paper. Also, k-means is known to be simple and computationally faster than other approaches such as graph clustering, and hierarchical agglomerative clustering approaches. In our implementation, we found that k-means works well in showcasing the database as a clean and rich repository from which to draw analytic conclusions.

We implement k-means in MATLAB using the squared Euclidean distance measure and the k-means++ algorithm for cluster center initialization.

Clustering via sequential k-means

While there has been considerable attention in increasing the optimality of a *streaming* variants of k-means[23][16], we are only interested in a solution that is motivated by k-means but works in a streaming mode. In comparison to the traditional batch/online methods, in sequential k-means, data points arrive in a stream of individual vectors. This algorithm is memory efficient as it retains only the i^{th} vector at any given time. Algorithm 1 presents the approach that we adopted. It starts with an initial k number of centers and refines the clusters with every new input vector. As the method tends to be noisy with random initialization, we look at our previously computed k-means cluster centers for initialization.

Algorithm 1: Sequential K-Means Algorithm

procedure SEQ K-MEANS$((x_1, x_2, \ldots, x_n), (m_1, m_2, \ldots, m_k))$ ▷ $k << n$
  Make inital guesses for the means $(m_1, m_2, \ldots, m_k)$
  Set counts of $(n_1, n_2, \ldots, n_k) \leftarrow 0$
  while *!feof* **do**
    Acquire next input, x ;
    if m_i *is closest to* x **then**
      $n_i \leftarrow n_i + 1$;
      $m_i \leftarrow m_i + (1/n_i) * (x - m_i)$;
    endif
  endwhile

Experimental Observations

We present the results of experiments comparing the different dancer datasets with k-means and then continue with the results of our implementation of the sequential k-means algorithm. Choosing the number of clusters in k-means is an unsolved problem though there has been considerable efforts in doing so [20]. We found that, after certain well-educated guesses, the result of the algorithm in our dataset did not vary much. Thus, by experiment, we settled on the number of clusters. We use a silhouette plot [15] to determine how well-separated the clusters are. The silhouette plot displayed below in Figure 1 is measure of how close each point in one cluster is to points in the neighboring clusters. We used this plot to experiment and validate the number of clusters used.

The results shown in Figure 2 and Figure 3 below compare the fingerprint of two dancers from backgrounds of Bharathanatyam and Hip-hop in particular. A similar prompt of 'space' [constraint] was provided to both the dancers.
An important observation that was made in the process of this analysis was the difference between the cluster center in the clusters and the data vector closest to it. Euclidean distance was used to find the vector closest to the average of the biggest cluster; for each run of the algorithm a study of the difference of the average to that vector was done. It was found that they were very similar, showing that the centroids represent an actual instance of the dance itself.

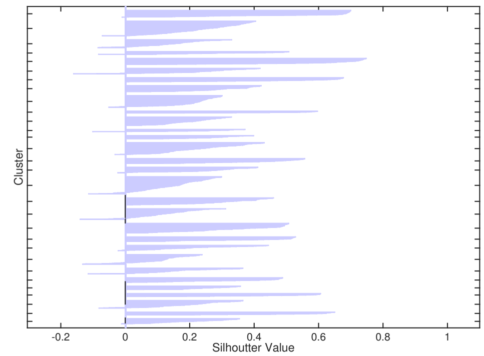

Figure 1. Silhouette plot showing 30 clusters.

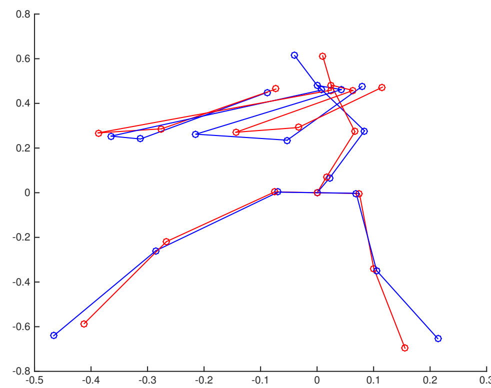

Figure 2. Habit extracted from the dataset of a dancer whose vernacular root is Bharathanatyam, an Indian classical dance. The red skeleton? shows the cluster center of the dominant cluster whereas the blue is the input vector closest to it.

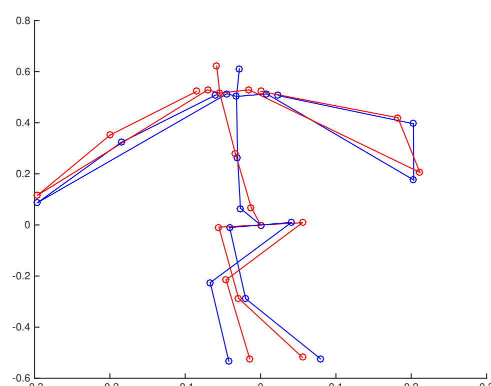

Figure 3. Habit extracted from the dataset of a dancer whose vernacular root is Hip-hop.

Figure 4 and Figure 5 depict the results from sequential k-means. The figures show four most dominant cluster centers in different iterations. These are extracted from the dataset of the dancer whose vernacular root is Bharathanatyam.

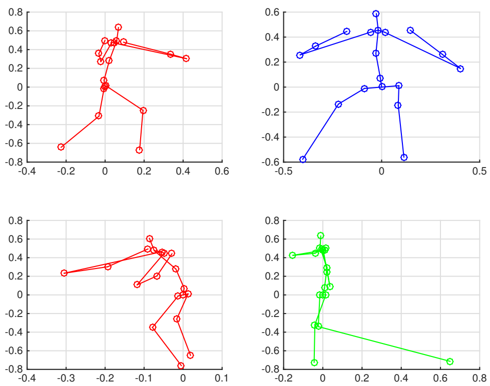

Figure 4. Four most dominant cluster centers in order of dominance.

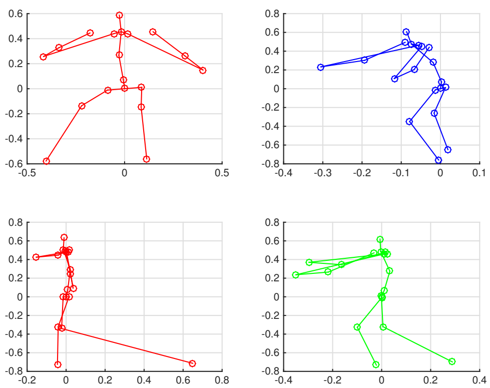

Figure 5. Four most dominant cluster centers in order of dominance in the 854th iteration.

The results clearly show the differences in the fingerprints of different dancers in their native and improvisational dance forms.

ONLINE FRAMEWORK

In designing the online community framework allowing access to the data, and upload of reinterpretation and annotation of the data, we look to recent examples such as Kickstarter, Github, and Stackoverflow. Each of these examples is an attempt to create a community "marketplace". On Kickstarter, groups pitch ideas in order to attract funding and publicity for their ideas. Github allows coders to create repositories of open source code which can be "branched" by other teams of coders and taken in new directions. Stackoverflow allows coders to ask questions and pose technical challenges, which are answered by a large community of experienced coders, where members curate the answers themselves in order to identify the most relevant and useful responses.

The design of our system is guided by a few simple needs (refer to Figure 6).

First, the Mo-Cap and video data must not only be made available, but also made as accessible as possible. This means

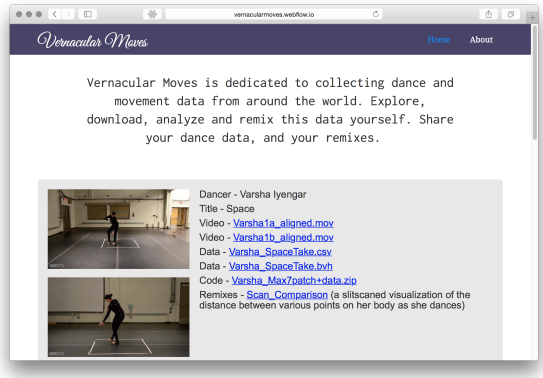

Figure 6. Online repository : Vernacular Moves

providing clean data in as many formats as possible, but also providing examples of how this data can be accessed, imported and manipulated using a variety of tools. Example code will be provided for working with the data in Max/MSP, MATLAB, Processing, X-Code, Maya, and Blender.

Next, once users have created new materials from materials in the database, there needs to be simple publishing tools to upload these responses including 3D files, images, videos, text and audio.

And finally, there will be some form of community critique and peer review system designed to allow the community to "self curate" incoming responses.

CONCLUSION AND FUTURE WORK

We do acknowledge and appreciate the contradictions in trying to capture the vernacular movement using digital/industrial tools, as well as the physical limitations these tools impose. Creating a mobile version of our motion capture rig is not feasible yet, making fieldwork impossible. Likewise we are not expecting that crowd-sourcing will immediately result in a creative community. This work represents an envisioning of a potential future world in which this library of motion portraits serves as the focal point around which a variety of stakeholders might form a new community of practice. This project begins as building a library or repository, yet our outcomes lie in how artists, designers, and scientists operationalize this database. Next steps for the project include a period of prototyping with the help of invited scholars and practitioners to test several models; not only for capturing and representing movement, but for community participation and critique of the works that grow from the source recordings. With people from a variety of backgrounds potentially interested in using this database to create new works, the growth and structure of the online community becomes a key component of this research. We look to engender distinctions between ways of analyzing and using the data that will be as varied, subtle, and dramatic as the differences in dance styles themselves, and work to reveal various "data cultures" at work in these fields.

ACKNOWLEDGMENTS

We thank Assegid Kidane, Qiao Wang, Shaun Ylatupa-Mcwhorter and the faculty, staff and students of the School of Arts, Media and Engineering.

REFERENCES

1. Repovizz: A multimodal on-line database and browsing tool for music performance research., 2010.
2. The Ohio State University. ACCAD - Motion Capture Lab - Data and Downloads. "Open Motion Data Project". `http://accad.osu.edu/research/mocap/mocap_data.htm`.
3. Barbič, J., Safonova, A., Pan, J.-Y., Faloutsos, C., Hodgins, J. K., and Pollard, N. S. Segmenting motion capture data into distinct behaviors. In *Proceedings of Graphics Interface 2004*, Canadian Human-Computer Communications Society (2004), 185–194.
4. Buchanan, P. A., and Ulrich, B. D. The feldenkrais method®: A dynamic approach to changing motor behavior. *Research Quarterly for Exercise and Sport 72*, 4 (2001), 315–323.
5. CMU Graphics Lab Motion Capture Database. `http://mocap.cs.cmu.edu/`.
6. Descartes, R. *Discourse on method, optics, geometry, and meteorology*. Hackett Publishing, 2001.
7. Duda, R. O., Hart, P. E., and Stork, D. G. *Pattern Classification*, 2 ed. Wiley Interscience, 2000.
8. Feldenkrais, M. *Awareness through movement*. Feldenkrais Resources, 1972.
9. Guest, A. H. *Kinetography Laban : the System of Analyzing and Recording Movement*. Theatre Arts Books, 1970.
10. Adrien M, Claire B, "Hakanai". http://www.am-cb.net/projets/hakanai/.
11. Illich, I. Vernacular values. *Philosophica 26* (1980).
12. John McCormick, Steph Hutchison, K. V. J. B. V. Emergent behaviour: Learning from an artificially intelligent performing software agent. *Proceedings of the 19th International Symposium on Electronic Art, ISEA2015* (2015).
13. Jules Françoise, Baptiste Caramiaux, F. B. Realtime segmentation and recognition of gestures using hierarchical markov models. In *Mémoire de Master, Université Pierre et Marie Curie–Ircam* (2011).
14. Kaschak, M. P., and Maner, J. K. Embodiment, evolution, and social cognition: An integrative framework. *European Journal of Social Psychology 39*, 7 (2009), 1236–1244.
15. L., K., and Rousseeuw, P. J. Finding groups in data: An introduction to cluster analysis. *Hoboken, NJ: John Wiley and Sons* (1990).
16. Liadan O'callaghan, Adam Meyerson, R. M. N. M. S. G. Steaming data algortihms for high quality clustering. In *IEEE Transactions on Knowledge and Data Engineering*, IEEE (2002), 0685.
17. Manning, S. *TDR: The Drama Review*. 2015.
18. Evan Roth, Ben Engebreth. http://whiteglovetracking.com/.
19. Motion Bank. http://motionbank.org/de.
20. Pham, D. T., Dimov, S. S., and Nguyen, C. Selection of k in k-means clustering. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science 219*, 1 (2005), 103–119.
21. Rushil Anirudh, Ahnaf Masroor, P. T. An online sampling for streaming video summarization.
22. Smigel, L., and Goldstein, M. *Documenting Dance, a Practical Guide*. Dance Heritage Coalition, 2006.
23. Sudipto Guha, Adam Meyerson, N. M. R. M. L. O. Clustering data streams: Theory and practice. In *IEEE Transactions on Knowledge and Data Engineering*, IEEE (2003/5), 515–528.
24. Wang, Q., Turaga, P., Coleman, G., and Ingalls, T. Somatech: an exploratory interface for altering movement habits. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, ACM (2014), 1765–1770.
25. Zhou, F., De la Torre, F., and Hodgins, J. K. Hierarchical aligned cluster analysis for temporal clustering of human motion. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 35*, 3 (2013), 582–596.