Analysis of Habitual Patterns in Vernacular Movement

by

Varsha Iyengar

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved April 2016 by the
Graduate Supervisory Committee:

Xin Wei Sha, Chair
Pavan Turaga
Grisha Coleman

ARIZONA STATE UNIVERSITY

May 2016

ABSTRACT

This thesis aims to explore the language of different bodies in the field of dance by analyzing the habitual patterns of dancers from different backgrounds and vernaculars. Contextually, the term habitual patterns is defined as the postures or poses that tend to re-appear, often unintentionally, as the dancer performs improvisational dance. The focus lies in exposing the movement vocabulary of a dancer to reveal his/her unique fingerprint.

The proposed approach for uncovering these movement patterns is to use a clustering technique; mainly k-means. In addition to a static method of analysis, this paper uses an online method of clustering using a streaming variant of k-means that integrates into the flow of components that can be used in a real-time interactive dance performance. The computational system is trained by the dancer to discover identifying patterns and therefore it enables a feedback loop resulting in a rich exchange between dancer and machine. This can help break a dancer's tendency to create similar postures, explore larger kinespheric space and invent movement beyond their current capabilities.

This paper describes a project that distinguishes itself in that it uses a custom database that is curated for the purpose of highlighting the similarities and differences between various movement forms. It puts particular emphasis on the process of choosing source movement qualitatively, before the technological capture process begins.

ACKNOWLEDGMENTS

I am indebted to various individuals who have provided me with crucial support and insight thus enabling the completion of this work.

My parents, the reason for the rich amassing experiences I hold so dear - I owe everything to them. My grandparents, whose timely calls and belief in me made all the difference in the world.

Dr. Sha Xin Wei, my thesis chair, who constantly broadened my perspective and let me dream of the many possibilities a practice could hold. For being such an approachable department head.

Dr. Pavan Turaga, who helped turn all the cloudy ideas into grounded realities. He made possible this entire work with an ever encouraging smile.

Prof. Grisha Coleman, who really took me under her wing. From talks of vernaculars to the different kinds of tea, she nurtured and made strong the ideas of somatics in me. She led me to the Arts, Media and Engineering Department, for which I will be forever grateful.

My colleagues, especially Rushil Anirudh and Qiao Wang whose help and advice greased the wheels; the faculty of Arts, Media + Engineering and the Computer Science department for imparting their knowledge and in particular Prof. Todd Ingalls for helping me lay a strong foundation in the department.

There are too many of my peers and friends who have shown unconditional love and encouragement. Akshay and Vincent who indulged in my wild theories while constantly challenging and questioning my thinking; Michael and Cooper, with whom I baked bread and explored the various dimensions of artistic work; Prashanth, a true comrade in this strange land; and finally my roommates, Arundhati, Kruthika and Devika who kept me well fed and had to put up with my whims and my dear friend, Adithya, who apart from being my 'coding-guru', helped remove the fear of unknowns and constantly made me believe in myself.

I am grateful to Ira Fulton School of Engineering, in particular, the School of Computing, Informatics, and Decision Systems Engineering, my parent college and the department of Arts, Media and Engineering for the financial support and where I can positively say that I found my true calling.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 Motivation

The integration of dance and technology has strong performative backing. Mercé Cunningham is famous for his explorations of technology – from film to video to computers. Alongside George Balanchine and Martha Graham, he is considered one of the most innovative dance choreographers of the century. In 1999, Cunningham created his famous work, *Biped* with Kaiser and Eshkar [13] where Cunningham dancers, wearing strategically placed optical sensors, were recorded in a motion capture system. The performance included ephemeral virtual dancers performing movement derived from the dancers as shown in Figure 1.1. In another striking work - Trisha Brown worked with a projector on her back.

These works indicate two parts of a whole – the dancer is linked with representations of his/her movement in a "coupled system". The notion of coupled systems has an anthropological account; Clark and Chalmers in *The Extended Mind* [11] state that -

"In certain conditions, the human organism is linked with an external entity in a two-way interaction. Creating a coupled system that can be seen as a cognitive system in its own right."

Although we do not delve into the vast field of cognition, we situate our work in the realm of dancer/machine interaction. While human movement became popular in the computer vision community, the approach to these relationships shifted from being a control relationship to a more collaborative one. Powerful motion capture technology like the OptiTrack Motive™ motion capture system, enables the capture and representation of complex human movement with a density of details. While methods of representation and classifi-

Figure 1.1: Cunningham, Biped, 1990

cation advance, our aim is enhance the performer/machine *collaborative* relationship. We base our ideas on the works of the choreographers and dancers mentioned above and bring these advances of technologies into play. As such, this project explores how movement identities can be an emergent property of data analysis.

As humans, our capacity to recognize and distinguish amongst different kinds of movement (animal, human, animate, inanimate) is a foundational evolutionary ability [22]. 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 motion capture system provides, it is clearly recognizable that sets of human movement data reveal certain unique and significant patterns. While there are resources that focus on quotidian, gestural movement and short sequences of movement, the basis of this research is on human motion data with a certain rooted vernacular context. In his 1980's book, *Vernacular Values*

[19], Ivan Illich defines vernacular work as work done outside any recognizable market in formal economics. This conceptual framework can be extended to the movement arts that has been resistant to methods of commodification, remaining in the realm of specialized forms of written movement notation for example, the Laban notation [17]. Through this, the work expands into native styles that are self-determined. By curating the data collection process to include varied vernacular movement, this project differentiates itself and caters to forms of non-European vernaculars that have defied the written notation systems for a host technical and cultural reasons. By recognizing the story of experience behind a dancer's movement form and vocabulary, emergent patterns in the movement vocabulary of the dancer will be easily distinguishable. Touching upon the subject of experience, Elizabeth A. Behnke, on detailing Edmund Husserl's phenomenology of embodiment [8], states that

"In Husserl's phenomenology of embodiment, then, the lived body is a lived center of experience, and both its movement capabilities and its distinctive register of sensations play a key role in his account of how we encounter other embodied agents in the shared space of a coherent and ever-explorable world."

Breaking the above statement down and re-connecting it to the process of discovering movement patterns in dance specifically in the context of performance, has profound implications. The "lived body" in question is the dancer that forms experiences of dancing - the internal experience as well as the two way interaction with the system and with the skeletal representation. Internal thoughts, sensations, information flow and other interactions form factors of said experience and these shape the identity of a person. Through this, the project questions the choice of movement of a dancer, not just the physical capability of the body. These questions form the philosophical context of the system. While the term "agents" could be misleading, the primary purpose of the project and the skeletal representations are to bring an awareness to the movements and thus possibly a change in the habitual choreographic patterns of an individual dancer.

At its core, this research focuses on understanding the statistical distributions that constitute an individual's motion data, magnifying the commonalities and differences that constitute habitual patterns, thereby exposing the unique vernacular of the mover. Detecting patterns in dance movement is key. In the context of creating interactive systems, the patterns allow the dancer to have a deeper influence on the computational system. By excavating into the movement vocabulary of the dancer, the system discovers the postures that are recurring frequently. This informs the dancer and can be used to bring self-awareness of one's dancing habits. Often while training, dancers are required to venture out of their comfort zone and break out of the rut of choreography. Projects like the Pathfinder [2] are developed to generate graphical patterns to help stimulate dancer choreography creativity. While the Pathfinder is a generative visual language that aims to inspire the dancer, by informing the dancer of their native body vocabulary, this work has a more direct approach to help the dancer break out of their artistic habits. Thus while improvising and creating movements, the dancer can break out of their tendency to create similar postures, explore larger kinespheric space and invent movement beyond their trained capability. This could lead to new inspiring work and less predictability and artist frustration.

In addition to the analysis, the thesis incorporates a new dimension of performance installation through its real time realization. MoCap is a state of the art system that is conventionally used for choreographed, synchronized full body recording. However, tracking and providing real time feedback through performance extends its capabilities and integrates it into the artistic component itself. Although designing the interaction is at its nascent stages, the visualization of the patterns on the fly forms an experiential system that was used to validate the experience.

1.2 Contributions and Formal Organization

To summarize the contributions of the thesis:

1. Creation of a motion capture database focused on the human movement data of different dance vernaculars.

2. Using methods of statistical data-analysis, mainly k-means clustering to discover the unique fingerprint of a movement data set.

3. Creation of an online pipeline that uses sequential k-means for a real-time interactive performance installation.

4. Enabling the extension of the capabilities of the available (to the research) motion capture system to include real-time data streaming.

5. Making available the code for the pre-processing of streaming motion capture data and sequential k-means on GitHub.

This entire document is divided into five chapters. The following chapter summarizes the work in the fields of somatic practices, motion capture datasets and movement analysis while drawing parallels to this work. By drawing on the works of practitioners in these fields, it provides context to this thesis and shapes the need for the development of such work. Chapter three introduces the proposed system and provides a detailed account of the pipeline of the project along with the hardware specifications of the tools used. The features extracted and the algorithms used to analyze them are discussed. Chapter four provides a discussion of the experimental observations and results. While the focus is on recognizing patterns in human motion, the system has an artistic component in that it provides a method to break out of choreographic patterns thus building an experiential interactive system for movement practitioners. Chapter five details an outcome of the thesis - a collaborative art installation called ‘Framed’ and the last chapter concludes by a discussion on the future scope of this thesis.

Chapter 2

BACKGROUND

2.1 Habitual Patterns

The human's evolutionary ability to learn sets us apart from other animals in different parameters [10]. This deep unconscious learning is informed by a variety of factors, including our physical and mental environments, our experiences and any other physical training such as sports or dance, that we undergo. The aim is to understand and discover this unconscious learning in terms of the movement arts.

This paper acknowledges that the notion and definition of habitual patterns is varied and exists in informatics, biotechnology and various other fields. Though these are often overlapping in nature, we are interested in its definition in the realm of somatics. Somatics is the field of movement studies with an emphasis on internal mindful sensations and action. Though the history and world of somatics is vast, we limit ourselves to the general idea of habitual patterns as defined by well-known somatic practitioners that use somatics as a method of therapy. The Alexander Technique and Feldenkrais Method are somatic educational systems that aim to reduce chronic pain by bringing awareness and eliminating unsound movement habits. Why We're in Pain [1], a book by Dr. Sarah Warren St. Pierre, a Certified Clinical Somatic Educator and co-owner of Somatic Movement Center, illustrates the notion of developing habitual patterns in the below quote.

".... *With each repetition, the movement pattern becomes more deeply learned. When our nervous system notices that we keep repeating the same movement or posture, it begins to make that movement or posture automatic. ...This process allows the parts of brain responsible for making voluntary decisions to focus on new things which require conscious*

attention."

While these texts are focused on everyday quotidian movement patterns, we can leverage these ideas to include patterns in dance movement. The movement vocabulary of a dancer is formed by various socio-economic, geographical and technical factors and is also largely attributed to the training in certain dance forms and styles he/she acquires over her lifetime. There are also several undercurrents of influences that affect a dancers movement style. It is understandable then that with sparse sets of data, these patterns can be easily discovered and used to compare vernaculars and through this, identify a movement sequence and possible a dancer. The results of the thesis is informed by such influences.

Narrowing it down, this work defines habitual patterns as postures that tend to re-appear as the dancer performs improvisational dance movement. This definition draws its basis from the idea of a certain style of movement being ingrained into a dancer's system during the course of his/her life. The goal of the system is to help break the dancer's tendency to create similar postures by bringing a certain external awareness of these postures. Moreover, branching away from the somatic idea of habitual patterns where there is a hazy distinction between nervous system and dancer, the process of the computational system is designed to learn the patterns and feed it visually to the dancer to enable change.

2.2 Prevalent Motion Capture Datasets

There are currently several open source libraries for motion capture data. Most of them focus of simple quotidian gestures like walking, running and jumping. Datasets such as the "Open Motion Data Project" at ACCAD [4] also make available certain dance data sets termed such as "BreakDance1" but these are limited at most. Possibly the most well known among these is the CMU Graphics Lab Motion Capture Database. It contains six categories ranging from physical activity and sports to human interaction with 23 subcategories of tracking data. The subcategory on 'dance' contains several ballet-based postures and motions with detailed description of the same. ‘GVVPerfCapEva’, by the Max Planck Institut Informatik, is a rich source of human shape and performance datasets that span different sensor modalities.

While these databases of human movement certainly exist, 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. It challenges the notion of digital neutrality and eventually aims to ask - *What might be revealed in recording, mining and comparing movement vernaculars?*

As other examples of datasets, there is a project by Evan Roth called “White Glove Tracking” [26] and a structurally similar collaboration between William Forsythe and AC-CAD “Synchronous Objects”. The results of these projects enrich the respective datasets by adding a layer of annotation through examples of choreography, data visualization and other sorts of re-interpretation by artists. Our aim is to curate the motion capture data set , not limited to a certain dance style, but to collect a wide range of vernacular movements in the dance field. The online repository that contains this data set is called "Vernacular Moves" [12]. Figure 2.1 provides a glimpse of one of the dataset available online.

This thesis leverages this unique dataset by adding an annotation layer by comparing the different movement vernaculars captured. This enhances the understanding of the movement style for a third-party user of the dataset.

2.3 Machine Learning in Movement Analysis

Aggarwal et al., provide a review of major areas related to the various tasks involved in interpreting human motion including motion analysis of human body structure, tracking of human motion with single and multiple cameras, and recognizing human activities from image sequences [5]. This paper is a result of the increasing attention human motion

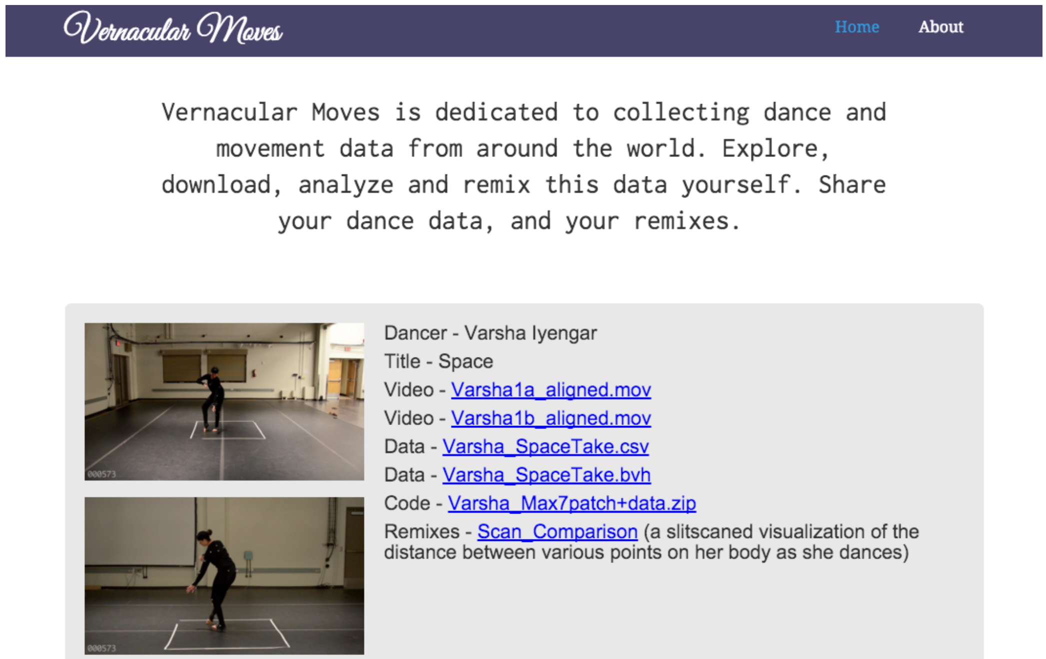


Figure 2.1: The Online Repository : Vernacular Moves

analysis is gathering from machine learning and computer vision researchers as well as the range of applications it caters to, such as athletic performance analysis and man-machine interfaces.

Possibly the most well explored among the various techniques in human motion recognition is gait analysis. Huang et al., demonstrate a method in their paper [18] that combines canonical space transformation and eigenspace transformation to recognize subjects by the way they walk. Results show that subjects can be recognized with an accuracy of 100 percent by this method. Nixon et al., proposed a system [27] that uses velocity, temporal symmetry and area masks in addition to silhouette-based analysis to describe the shape as well as how a particular subjects move. They also describe a method of signature extraction to recognize an individual. From these works, we derive the strong notion of a signature of movement that can be identified and even derived from each individual. Generally speaking, somatic education emphasizes creating conditions for more efficient, functional movement patterns to emerge [16] [34]. 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. The above works point out that analysis methods can extract movement patterns from a sample set of human movement databases to reveal distinct 'signatures'. Stevenage et al., in two experiments conducted [32], prove that gait signature was sophisticated enough to learn to identify six individuals under different and adverse lighting conditions. We are interested in exploring a broader range of differences, not in the gait, but in the movement of a dancer. The notion of identifying an individual with respect to their movement style remains the same even though the means to achieve it vary.

There are other examples of machine learning in movement analysis that depart from the standard analysis on short gestures or a specific movement. Barbic et al., proposed an algorithm to decompose human motion and places a cut when the distribution of human poses changes based on probabilistic principal component analysis (PCA) [7]. Comparing it to this thesis, our work does not delve into recognition or decomposition of gesture itself, but in visualizing dominant clusters in the movement sequence. Françoise et al., proposed a method to analyze movement sequences based on motion trajectory synthesis with Hidden Markov Models [21]. Their aim was to investigate the consistency, in terms of the repeatability of acceleration patterns and angular variations. They developed their experiments in the use case of two performers with a different level of expertise in Tai Chi. Several works draw upon various domain specific features such as models representing movement qualities in dance [6] or the segmentation and recognition of a set of movement primitives from larger action sequences [25]. The features used in this project are positional information of 19 joints forming a skeletal framework. Rather than recognizing patterns, the focus is on unsupervised learning via clustering to discover the unique fingerprints or postures in the data set.

In the context of interactive dance performances, detecting patterns in movement allows a stronger collaborative relationship between the dancer and the machine. These interactions can be seen in contemporary dance, art installations, and in club settings. Before diving into the question of using patterns in particular to enhance computational learning, referencing the work of John McCormick with *Emergence* [20] becomes imperative as it is the main source of inspiration for this work. Created by John McCormick and Steph Hutchison in collaboration with other organizations, Emergence uses Artificial Neural Networks and Self Organising Maps to identify and create a "shared vocabulary" with the dancer and a digital performing agent.

Figure 2.2: Emergence by John McCormick, Steph Hutchison And a Digital Performing Agent. [24]

The agent follows the dancer, learns and creates movement phrases as a result of the dancers captured data. While the vision of this work extends to creating an agent that improvises with the human dancer, the immediate goal is to identify the unique fingerprint of the dancer thereby feeding the computational system information unique to every body. We choose to use sequential k-means to cluster and find this fingerprint rather than getting lost in the multiple layers that constitute methods of data analytics such as Artificial Neural Networks.

There have been other works focused on detecting patterns in dance movement. Pohl et al. describe a method to detect patterns in dance movements using dynamic time warping (DTW) [30]. They used DTW to compute the similarity of two movement sequences that may differ in temporal domain. By running the DTW algorithm, they found two distances: a distance between the sequences, and a warp-path that describes the alignment of the sequences. They assumed no *apriori* knowledge on possible patterns and thus used a custom threshold unsupervised clustering algorithm to assign labels. This method resulted in error rates of about 20-30 percent in distinguishing movements. It was indicated that more pronounced movements could be classified better. Tang et al., [33] developed an algorithm to find repetitive dance patters in 3D motion capture data. Repetitive patterns were traced from a point cloud of similar postures via a similarity matrix. These patterns were traced through DTW as well. Auto-clustering and pattern alignment was used to classify patterns and an estimate of the cycle period was computed. Tang et al., also mention that, in a future work, these repetitive patterns can be used to summarize a piece of motion capture data and that they can be used as features that define a motion clip uniquely. Needless to say, by defining an entire motion clip, we can define the signature unique footprint from that clip and call it the habit of the dancer.

2.4 Real-time Approaches to Clustering

Clustering, in broad strokes, is the problem of partitioning the data set so that "similar" (defined in terms of the problem) items are in the same partition. This work focuses on the definition of clustering brought about by the k-means objective which is of identifying k centers so that the sum of distances from each point to the nearest cluster in minimized. Moreover, the paper studies the k-means objective in a streaming context.

Streaming data analysis has gained traction due to its lower memory consumption. A dataset is inconvenient to store when there is a large volume of data arriving continuously. A variety of approaches have been applied to real-time clustering. For diverse sampling of streaming video, Anirudh et al. use a generalization of the online K-means algorithm by modifying the cost that includes a penalty score if two cluster centers are too close in the feature space [31]. The results fare better than the standard online K-means clustering and the k-mediods clustering algorithms. Streaming motion tracking data poses tremendous challenges due to the size of the stream as well as the noise caused due to imperfections in the capture process. While most of the noise is removed by pre-processing the data, the method used needs to be simple and fast and provide results in a dimension understandable to the dancer. O'Callaghan et al., describe a streaming algorithm that effectively clusters large data streams [28]. The data arrives in chunks and the incoming chunk is clustered using a custom LSEARCH algorithm that is a subroutine to a k-Median algorithm and refines an initial solution with local improvements. The median is assigned a weight equal to the sum of the members and LSEARCH is re-applied to only the retained k weighted cluster centers. This process increases the cumulative running time of the clustering process but the quality of clustering was high. Guha et al., detail the need for a clustering stream model by comparing its advantages to the online and incremental models. They go on to define the k-means and k-mediod objective in the streaming context and provide a streaming algorithm that requires only a single pass. The algorithm presented is based on a facility location algorithm could produce more than k centers. The resulting algorithm has a polynomial run-time. Comparing it to the standard k-means, there was an expected trade-off between the cluster quality and the running time. The facility location-based algorithm produced clusters of near-optimum quality as compared to k-means that produced these solutions with inferior quality. The algorithm however took much longer to find better answers. Moreover, Chakraborty et al., compare the performance of the incremental k-means and incremental DBSCAN algorithm on a dynamic air pollution database [9]. The results of the paper conclude that the incremental k-means performs faster since it takes less time for the changes to reflect whereas DBSCAN needs the extra time to handle outliers in the data. As the prime intention of this work is for a real-time performance, the updation of the clusters to find better and more learned movement postures in a timely manner is more important than the highest quality of clustering.

Another method that can be used for sequential clustering is the self-organizing map (SOM). SOM is a type of Artificial Neural Network that can produce a discretized representation (map) of high-dimensional data. It consists of components called the nodes that are each associated with a weight vector. The procedure for placing a vector from data space onto the map is to find the node with the closest (smallest distance metric) weight vector to the data space vector. Although the updation of the weights are done in a different manner, this process can be modeled to sequentially compute the winner neuron and update the weights accordingly [1]. SOM algorithms are known to cluster small datasets efficiently and are widely known for pattern recognition [3]. Since this thesis is focused on discovering patterns in the dataset rather than classifying the patterns into particular styles of movement, a variant of the standard k-means was implemented.

The optimization of the standard k-means objective function is known to be NP-hard. Most of the work in the field thus focuses on the optimization of the algorithm as well as creating new algorithms whose objective function value is a fixed known value. The thesis does not aim to optimize k-means or present a solution that is highest in clustering quality, but in using an algorithm that can be run easily with different and varying movement datasets in a streaming real-time setting. In the interest of this simplicity and computational efficiency, we choose to use k-means and an online variant of k-means clustering inspired by Duda k-means [14], which allows processing of movement data as it arrives, and updates clusters on the fly.

Chapter 3

PROPOSED SYSTEM

3.1 Overview

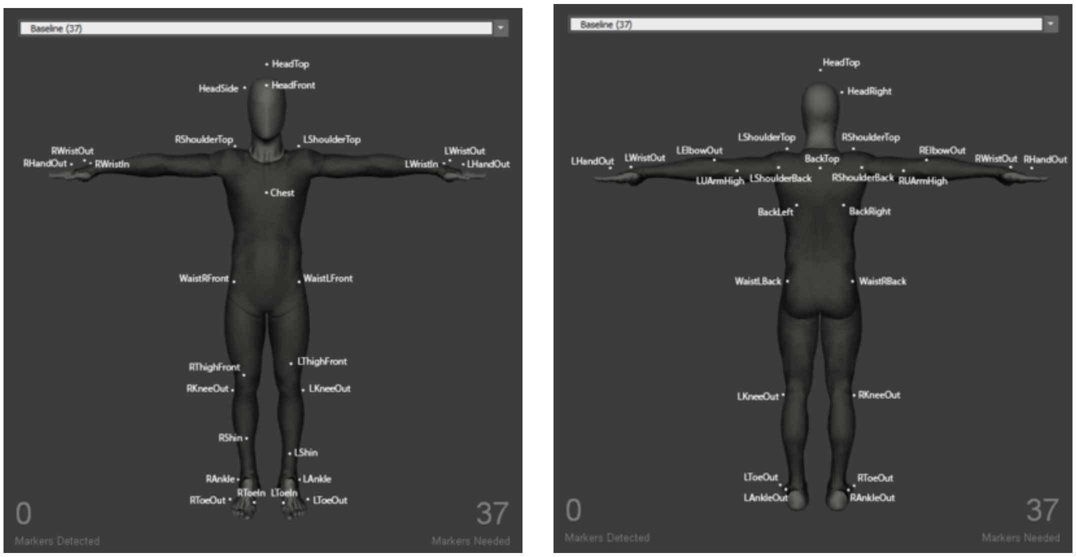
This thesis presents a system that draws several influences from the above mentioned works. With the purpose of a performative installation, the methods and technologies used to realize a real-time interactive system are detailed in this chapter.

3.2 System Hardware

For the motion capture sessions, two streams of data were captured in parallel including a primary stream consisting of optical motion capture using the OptiTrack Motion Capture system. We use a 12-camera system running at 120 FPS for high precision. OptiTrack's Motive software platform is designed to control motion capture system for rigid bodies as well as the full body skeleton.

Human movements are tracked through skeleton assets. We use the "Motive:Body" software to track a performer wearing a suit with 37 retro-reflective markers.

Motive uses pre-defined skeletal marker sets and the markers must be placed in these specific locations for a skeleton asset to be correctly recognized and configured. For a detailed description of the locations of the markers on the suit, refer to Figure 3.1a and Figure 3.1b. While tracking, 3D coordinates are re-constructed from 2D views seen from each camera in the system. Using these coordinates, we obtain six degrees of freedom - 3D position on the X, Y and Z planes as well as orientation data and this enables capture of complex movements in the capture space.



(a) Front view

(b) Back view

Figure 3.1: Baseline 37 Markerset

3.3 Methods of Data collection

For the sake of current analysis, movement data was captured from practitioners of different dance styles including African-American Diaspora, Hip-hop and Bharathanatyam. Simple prompts with parameters around axes of rhythm, space, and duration were created keeping in mind the diversity of the styles. 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. Making certain that these prompts were simple enough to not increase the cognitive load on the dancers, their sole purpose was to help in finding a common basis for further analysis and aiding the capture process. The final pose representation is simply a point-cloud of 3D locations of the joints, with the hip joint considered as the origin.

We take advantage of both the recording and exporting as well as the data streaming capabilities in Motive. For the initial analysis sessions, the recorded motion capture data was exported in the CSV format. For data streaming, we use the *NatNet SDK* that contains a native C++ networking library and the dynamic import libraries. *Open Sound Control (OSC)* was the protocol used for the communication between the Motive system running an *OSC Client* application in C++ and the system running a sample *OSC Server* application. The *OSC packet*, a unit of transmission of OSC, is represented by a datagram by the network protocol UDP. This way the joints forming a skeleton in each time frame was "bundled" into a message and transmitted as such to the OSC Server. An *OSC message* consists of an *OSC Address Patterns* followed by the *OSC Type Tag String* followed by a number of arguments. Our OSC message is of the format :

/skeleton <skeletonID> <boneID> <boneX> <boneY> <boneZ> ... <boneID> <boneX> <boneY> <boneZ>

3.4 Data Pre-processing

The data streamed from the Motive MoCap system has certain bone naming conventions and these are autolabeled as the streaming continues. Before analysis, the raw data needs to be cleaned to suit our purposes.

The first step is to handle missing data. Although we reduce losing data due to marker occlusions by correct positioning of markers and conditioning the movement, there are certain postures that result in erroneous data. These missing values are filled in with values of the joints from the previous frame. Since we are tracking continuous motion, we can rely on the assumption that the adjacent position of the joints will not be radically different from the previous one.

We consider only the positional data and eliminate the orientation data. It is thus straightforward to see the relative motion of a limb as a circle centered about the joint where the limb is attached. The resulting dataset has 19 joint positions of three dimensions each (X, Y and Z) totaling 57 dimensions. Locating the hip joint, we subtract the hip positions from the rest of the data. This makes the skeletal data invariant to the absolute location of the dancer in the real-world setting. It places the hip joint at the origin converting all the 3D joint coordinates from the world coordinate system to a person-centric system.

3.5 Clustering via K-means

A common form of data analysis involves clustering, a method of partitioning the data set into sets or clusters of similar items. 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 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. [14]

The objective is to find the *local* minima of:

$$\sum_{k=1}^{K} \sum_{i \in c_k} ||x_i - m_k||^2 \qquad \text{Euclidean distance}$$

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

Informally, the steps of the algorithms are as follows:

1. Start with initial guesses for cluster centers (centroids)

2. For each data point, find closest cluster center (partitioning step)

3. Replace each centroid by average of data points in its partition

4. Iterate until convergence

The most common algorithm [14] uses an iterative approach. Given the initial set of k means, $m_1, m_2, ...m_k$, the two important steps as listed in [23] in the algorithm are the :

1. **Assignment step :** Each incoming vector is assigned to the cluster whose mean yields the least within cluster sum of squares as iterated in the equation above. The sum of the squares is the squared Euclidean distance. The Euclidean distance formula is as follows:

$$\sqrt{\sum_{n}^{i=1}(q_i - p_i)^2}$$

Mathematically, it is partitioning the observations according to the Voronoi diagram and the following formula:

$$S_i^{(t)} = \left\{x_p : \left\|x_p - m_i^{(t)}\right\|^2 \leq \left\|x_p - m_j^{(t)}\right\|^2 \forall j, 1 \leq j \leq k\right\}$$

where each x_p is assigned to exactly one $S^{(t)}$ region.

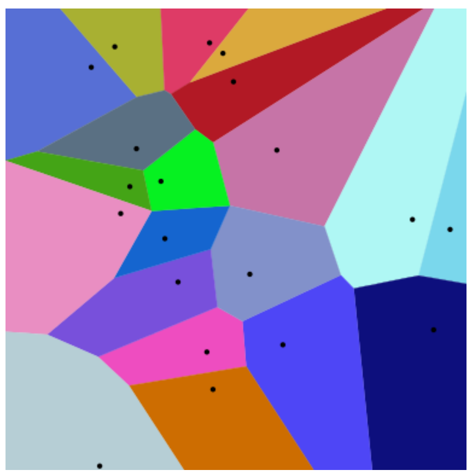


Figure 3.2: Euclidean Voronoi Diagram [15]

2. **Update step :** In this step, the new centroids are re-calculated in the new clusters.

$$m_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j$$

This also reduces the within-cluster sum of squares distance. The algorithm converges when the assignment no longer changes. Through observational familiarity, we choose the value of k in this way.

Our reasons for using k-means as the clustering technique is two-fold. It enables extending the approach to a real-time setting via variants such as sequential k-means. We elaborate on this method in the next section. Also, k-means is known to be simple and computationally faster than other approaches such as graph clusterings, and hierarchical agglomerative clustering approaches [].

3.6 Clustering via Sequential K-means

As commercial and scientific data sources continue to grow, it is imperative to use algorithms that are light-weight and can operate online, or in streaming settings. A data stream is an ordered sequence of incoming data vectors $x_1, x_2, \ldots, x_n$ that can be accessed only in the order it is received and can be read from the stream only once i.e at a single pass. we have noted earlier that motion tracking data is often tremendous in size and noisy due to the capture process. For our purposes, streaming data becomes important in two scenarios :

1. For a real time performance setting, the system has to work with any dancer/performer in space updating the clusters and identifying the habit on the fly.

2. Estimation of the time required for a dancer to learn and incorporate a pattern in their improvisational dance is unfeasible as it is dependent on the dancer.

3. This algorithm is memory efficient as it retains only the i^{th} vector at any given time.

While there has been considerable attention in increasing the optimality of *streaming* variants of k-means, 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 of clustering, in sequential k-means, data points arrive in a stream of individual vectors as explained above. 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. It updates the means one at a time rather than all at once. Hence, the clustering starts before all the vectors are acquired.

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($(x_1, x_2, \ldots, x_n), (m_1, m_2, \ldots, m_k)$) $\triangleright\ k << n$

  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

Chapter 4

EXPERIMENTAL OBSERVATIONS AND RESULTS

This chapter presents the results of experiments comparing the different dancer datasets with k-means and then continues with the results of the implementation of the sequential k-means algorithm.

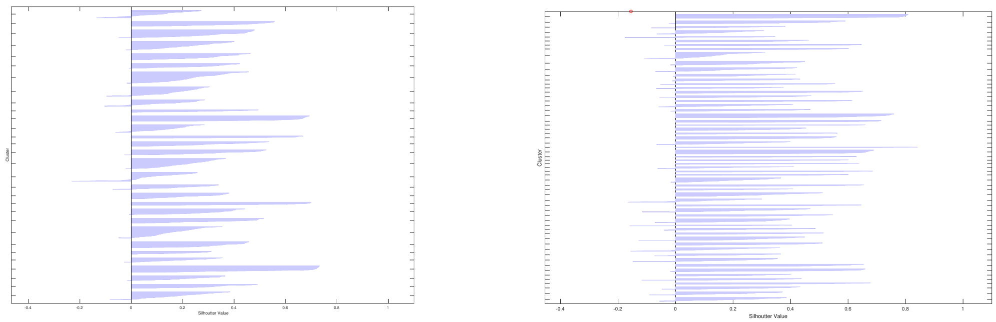
4.1 Choosing k

Choosing the number of clusters in k-means is an unsolved problem though there has been considerable efforts in doing so [29]. It was found that, after certain well-educated guesses, the result of the algorithm in the datasets did not vary much. Thus, by experiment, the number of clusters, i.e k was chosen.

To determine how well separated the clusters are, a silhouette plot is examined. This gives us another estimate to choose k. The silhouette value for each point, in this case vector, is a measure of how similar that vector is to others in its own cluster when compared to vectors in other clusters. Mathematically, it is defined by :

$$S_i = (b_i - a_i)/max(a_i, b_i)$$

where, a_i is the average distance from i^{th} data vector to others in the same cluster and b_i measures the distance to vectors in other clusters. We plot the silhouette cluster plot using the squared Euclidean distance. The measure ranges from +1, indicating well-separation of clusters to -1, indicating vectors that are probably assigned to the wrong cluster. The plots displayed below in Figure 4.1a, Figure 4.1 and Figure 4.2 show the experimental plots that determined separability of the clusters resulting also, in the number of clusters chosen.



(a) $k = 30$ (b) $k = 60$

Figure 4.1: Silhouette Plot Using Different Values of k

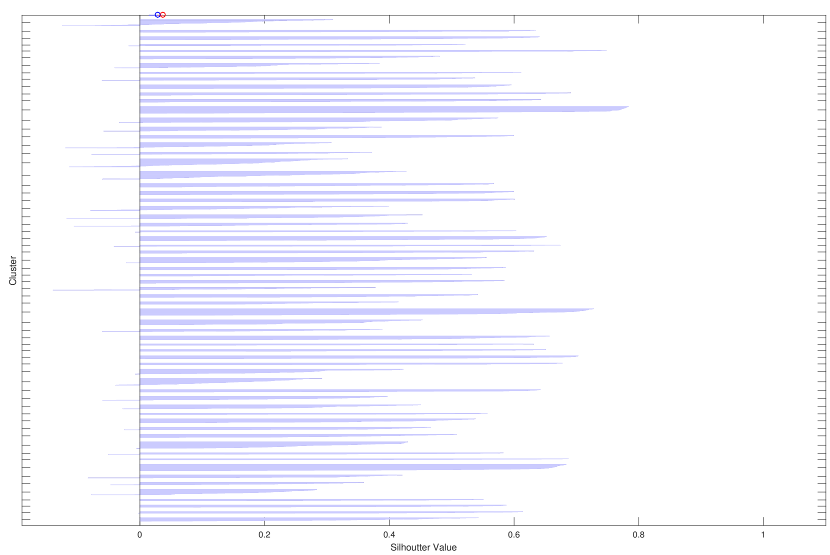
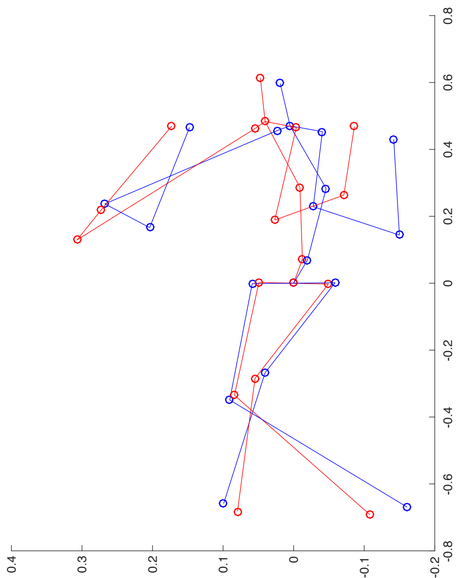


Figure 4.2: Silhouette Plot Using $k = 65$

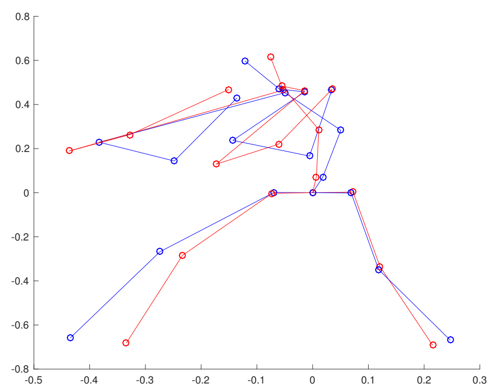
The results of experimentation with the dancer data whose vernacular root is Bharathanatyam, an Indian Classical dance form are displayed. Similar experimentation with different datasets resulted in a value of k for each different dataset. Since the number of negative clusters were lesser in the plot with 65 clusters, k was chosen as 65.

4.2 Discovering Vernaculars

Finally, the results shown in Figure 4.5 and Figure 4.4 below shows 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.



(a) YZ View



(b) XY View

Figure 4.3: Habitual Posture Extracted from the Data Set 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 Skeleton Is the Input Vector Closest to It.

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 i.e the red and blue skeletons in the above graphs. Euclidean distance was used to find the vector closest to the centroid of the dominant 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.

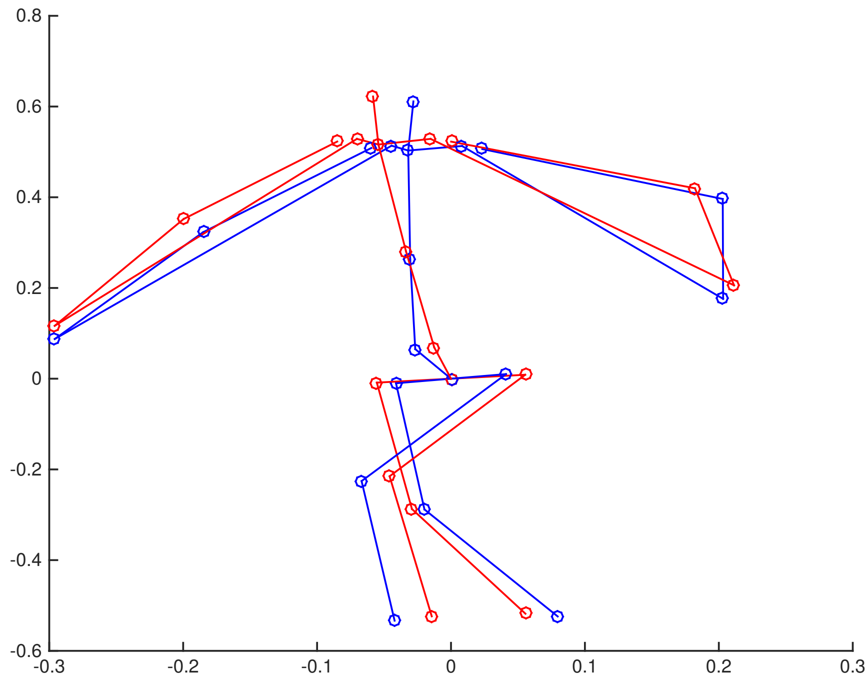
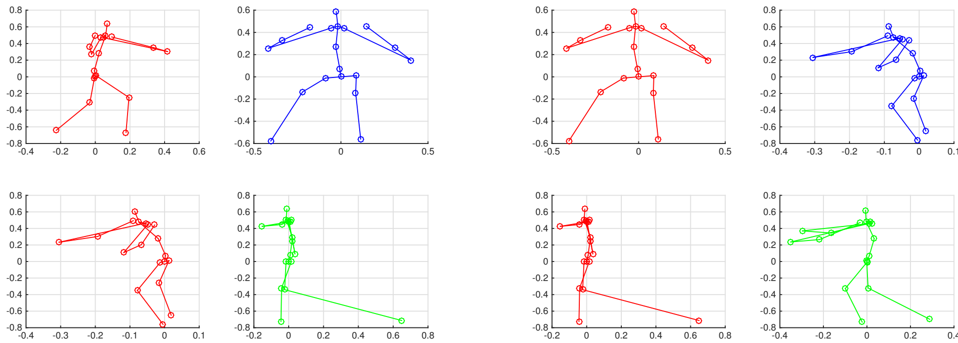


Figure 4.4: Habitual Posture Extracted from the Dataset of a Dancer Whose Vernacular Root Is Hip-hop.

4.3 Habitual Postures via Sequential K-means

This section depicts the results from the sequential k-means algorithm. Figure 4.5a and Figure 4.5b show the four most dominant cluster centers in different iterations of the dataset whose vernacular root is Bharathanatyam. These cluster centers, when displayed in such an iterative fashion informs the dancer of the various postures that tend to re-appear in an improvisational dance capture session. The habitual posture or fingerprint thus changes in real-time depending on the number of the clusters chosen and the dataset.



(a) 120th Iteration of the Algorithm. (b) 854th Iteration of the Algorithm

Figure 4.5: Four Most Dominant Cluster Centroids or Habitual Postures Discovered at Different Iterations.

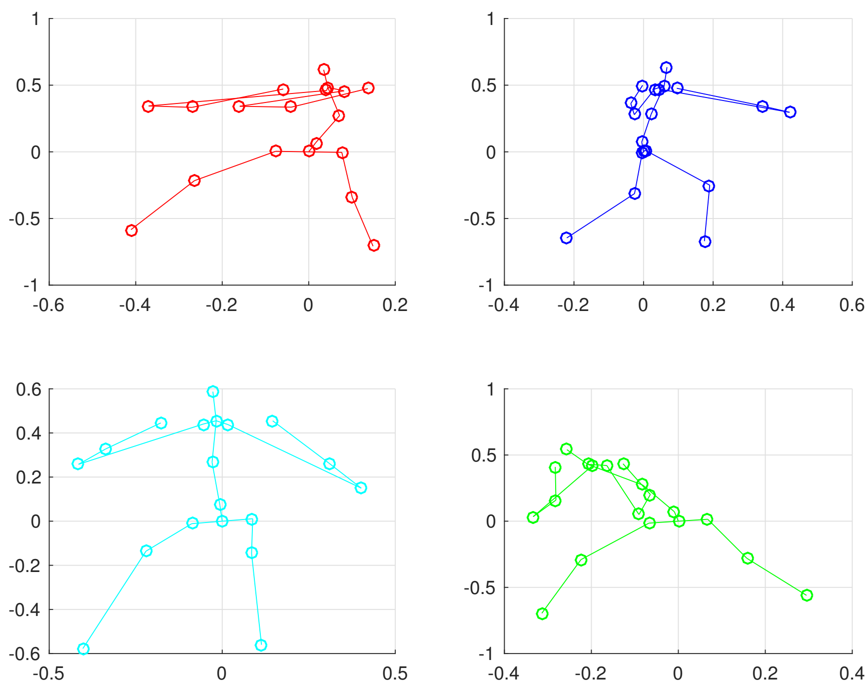


Figure 4.6: Habitual Postures Extracted from the Data Set Whose Vernacular Root Is Bharathanatyam.

As mentioned earlier, the sequential k-means algorithm is derived from the standard k-means clustering algorithm while differing in two notable aspects:

First, the cluster centroids are updated one at a time rather than all at once. This feature essentially enables us to start clustering while the dance motion capture is in progress. Second, this particular implementation of the algorithm uses pre-computed k-means centroids for initialization before streaming commences. It is therefore valuable to note that the iterations occurring as the capture proceeds show instances of "learning" in the algorithm. The final set of postures discovered converge and are similar to the postures obtained from the k-means algorithm for the same data set. We use the same set of centroids for the different data sets.

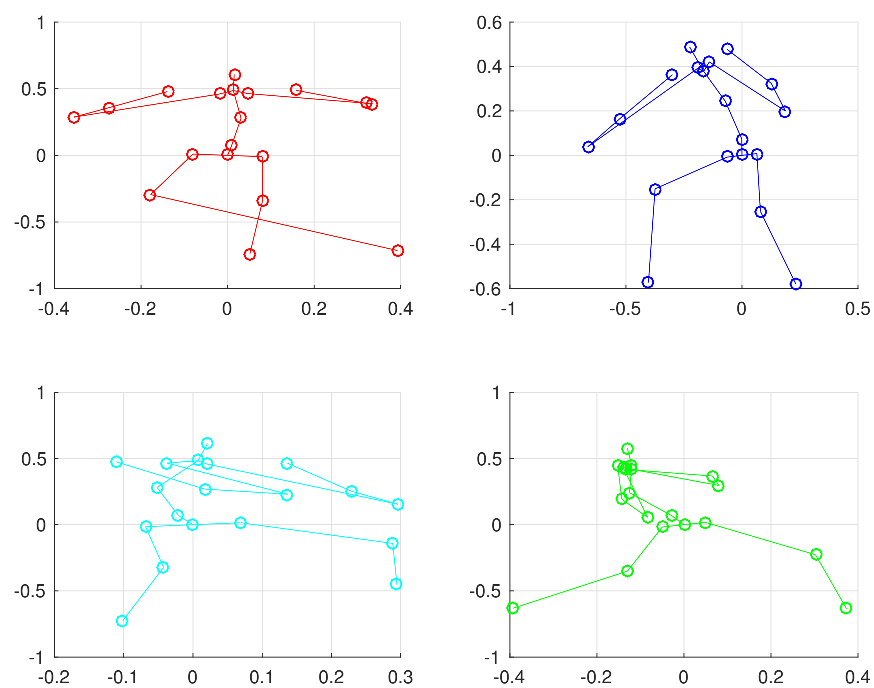


Figure 4.7: Habitual Postures Extracted from the Dataset of a Dancer Whose Vernacular Root Is Hip-hop.

The results clearly show the differences in the fingerprints of different dancers in their improvisational dance forms. It should also be noted that in the second data set, the a single dominant cluster was updated for a long time before all the clusters were updated.

4.4 Validation

Verifying the output of the algorithm with the actual instance of the performance forms a crucial validation task in the research. This helps identify the differences between the postures of the vernaculars. As there are no formal and practical methods of calculating the accuracy of the clusters in the k-means algorithm, user validation becomes imperative to determine the quality of the work as well as aids int eh process of self-reflection of the dancer.

It should be noted that the primary data set consists of the movement data of the author. As the vernacular root of the data set is known to be Bharathanatyam, by comparing the video of the dance during the capture and with existing knowledge of the dance form, the fingerprint derived from the system coincides with the movement style of the dancer. future work will include more neutral forms of validation.

Chapter 5

FRAMED : REAL-TIME DISCOVERY OF PATTERNS

The vision of this thesis is to create an interactive performance environment that results in a collaborative relationship between dancer and computational system to enhance the creative process of movement making. While *Emergence* created a digital agent that that learns and understands human movement, there is value in the human image as pointed out by *Biped*. By capturing as much information about a person , their story becomes visible and thus easily relatable for the dancer themselves to be informed of their movements. *Framed* is an attempt at creating a performance environment that is influenced by both these works creating a platform for self-awareness as well as an enhanced method for the system to understand the dancer's movement.

Framed leverages the postures identified by the clustering methods in this work. Working in collaboration with Jennifer Weiler, the wire-frame models of the postures were combined with *Mushi* [35], a generative visual art project to better represent, in terms of the visualizations, the skeletal forms. Mushi, created in Processing, is a self-designing artwork that is used to generate unique paintings. There are several iterations to this project; in the case of this project, the distortion is cause by black and white raster images of the postures. The postures are converted into black and white images that are then used as the source image to control the parts of the canvas being drawn in. The following image Figure 5.1 shows the result when only the lighter areas are drawn in.

This visualization forms the last block of the online pipeline. The postures produced from the sequential k-means algorithm are stored in real-time. As the algorithm learns the patterns of the dancer, the resulting posture changes and the current generated image overwrites the previous result. Thus, the code in Processing is constantly fed with a new

Figure 5.1: Visualizing Habitual Patterns with Mushi.

image resulting in continuous visualizations of the recurring patterns.

While running and testing the entire system in a real-time performance environment remains in the scope of future work, a static version of Framed was showcased in the Digital Culture Showcase in December 2015 at Arizona State University. This was an art installation where the unique postures or fingerprints derived from different dancer datasets were projected into wooden frames. A simple module in Max/MSP was used for the purpose of projection mapping into the frames.

The installation at the DC Showcase was well received. The main feedback pertained to the visualization of the skeletons. Future work thus includes a better mapping between the generative brush strokes to the movement itself and not just the still image.

Chapter 6

CONCLUSION

This thesis presents a new perspective on data collection and analysis in the context of interactive dance performance. Through the process of curating a motion capture database, it explores various non-European vernaculars that resist realized forms of notation due to various cultural and technical reasons.

It leverages this database by excavating into the movement vocabulary of the different dancers in this varied data set to discover the unique fingerprint that is the signature of the data set. The results of this is used to add a layer of annotation to the dataset to provide a deeper understanding of the movement to a third-party user. By discovering these postures, we make comparisons between movement signatures of different dancers and possibly extend that to the movement styles themselves. To enable this we use data analytics tool such as k-means clustering.

We also propose a real time solution in the form of sequential k-means. In building a complete plug and play framework module, this work allows experimentation with other features vectors of the dataset and other use cases.

Ultimately, this flow of components culminates in the making of a collaborative performance environment. The aim is to branch away from performances where the dance and visual effects are thought of as separate entities. The visuals and the dancer in this work create a feedback loop that is beneficial to the dancer and is designed to be aesthetically enhanced for an audience. The vision is to create a gallery of postures the dancer is situated in and enable all these postures to contribute to exploring and creating movement that is not native to the body.

Chapter 7

FUTURE WORK

The immediate future work is to validate the patterns discovered in a neutral manner and to test the system on different dancers by expanding the database further to include a larger set of movement vernaculars. Designing a more enhanced interaction between dancer and machine forms an important step towards creating a tool dancers can use to help explore a wider choreographic space. Feedback from dancers will be used to create he visualization of the postures in terms of the time taken for the postures to change, the colors and other elements of design.

Since the data is streamed in real-time from the motion capture system, the marker occlusions cause the data to be noisy. Handling missing data and testing in real-time to create a robust system was the biggest challenge and this needs to be tuned more efficiently for future usage.

The entire work-flow and code, including the method of cleaning the motion capture data along with the implementation of the k-means and sequential k-means algorithm will be available on Github.

REFERENCES

[1] Computer Vision Lab, Dresden, 1 edition. URL `http://www1.inf.tu-dresden.de/~ds24/lehre/ml_ws_2013/ml_09_clust.pdf`.

[2] 2015. URL `https://vimeo.com/111361109`.

[3] O. A. Abbas et al. Comparisons between data clustering algorithms.

[4] ACCAD. The Ohio State University. ACCAD - Motion Capture Lab - Data and Downloads. "Open Motion Data Project". 2003. `http://accad.osu.edu/research/mocap/mocap_data.htm`.

[5] J. K. Aggarwal and Q. Cai. Human motion analysis: A review. In *Nonrigid and Articulated Motion Workshop, 1997. Proceedings., IEEE*, pages 90–102. IEEE, 1997.

[6] S. F. Alaoui, B. Caramiaux, M. Serrano, and F. Bevilacqua. Movement qualities as interaction modality. In *Proceedings of the Designing Interactive Systems Conference*, pages 761–769. ACM, 2012.

[7] J. Barbič, A. Safonova, J.-Y. Pan, C. Faloutsos, J. K. Hodgins, and N. S. Pollard. Segmenting motion capture data into distinct behaviors. In *Proceedings of Graphics Interface 2004*, pages 185–194. Canadian Human-Computer Communications Society, 2004.

[8] E. A. Behnke. Husserl's phenomenology of embodiment. In J. Fieser and B. Dowden, editors, *Internet Encyclopedia of Philosophy*. 2011.

[9] S. Chakraborty, N. Nagwani, and L. Dey. Performance comparison of incremental k-means and incremental dbscan algorithms. *arXiv preprint arXiv:1406.4751*, 2014.

[10] D. J. Chalmers. The evolution of learning: An experiment in genetic connectionism.

[11] A. Clark and D. Chalmers. The extended mind. *analysis*, 58(1):7–19, 1998.

[12] G. Coleman, D. Tinapple, V. Iyengar, and S. Sagheb. Vernacular moves, 2015. URL `http://vernacularmoves.webflow.io/about`.

[13] A. Dils. The ghost in the machine: Merce cunningham and bill t. jones. *PAJ: A Journal of Performance and Art*, 24(1):94–104, 2002.

[14] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley Interscience, 2 edition, 2000. URL `/bib/duda/Duda2000/DudaR.O.,HartP.E.,StorkD.G.Patternclassification(2ed.,Wiley)(C)(738s).djvu`.

[15] B. Ertl. *Euclidean Voronoi Diagram*. 2015. URL `https://commons.wikimedia.org/wiki/File:Euclidean_Voronoi_diagram.svg`.

[16] M. Feldenkrais. *Awareness through movement*. Feldenkrais Resources, 1972.

[17] A. H. Guest. *Kinetography Laban : the System of Analyzing and Recording Movement*. Theatre Arts Books, 1970.

[18] P. S. Huang, C. J. Harris, and M. S. Nixon. Recognising humans by gait via parametric canonical space. *Artificial Intelligence in Engineering*, 13(4):359–366, 1999.

[19] I. Illich. Vernacular values. *Philosophica*, 26, 1980.

[20] K. V. J. B. V. John McCormick, Steph Hutchison. Emergent behaviour: Learning from an artificially intelligent performing software agent. *Proceedings of the 19th International Symposium on Electronic Art, ISEA2015*, 2015.

[21] F. B. Jules Françoise, Baptiste Caramiaux. Realtime segmentation and recognition of gestures using hierarchical markov models. In *Mémoire de Master, Université Pierre et Marie Curie–Ircam*, 2011.

[22] M. P. Kaschak and J. K. Maner. Embodiment, evolution, and social cognition: An integrative framework. *European Journal of Social Psychology*, 39(7):1236–1244, 2009. ISSN 1099-0992. doi: 10.1002/ejsp.664. URL `http://dx.doi.org/10.1002/ejsp.664`.

[23] D. MacKay. An example inference task: clustering. *Information theory, inference and learning algorithms*, pages 284–292, 2003.

[24] J. McCormick. *Emergence by John McCormick, Steph Hutchison and an emerging performing agent, is a duet performed between a human dancer and an artificially intelligent performing agent.* URL `http://www.johnmccormick.info/emergence/`.

[25] F. Meier, E. Theodorou, and S. Schaal. Movement segmentation and recognition for imitation learning. In *International Conference on Artificial Intelligence and Statistics*, pages 761–769, 2012.

[26] motionbank. Evan Roth, Ben Engebreth. http://whiteglovetracking.com/.

[27] M. S. Nixon and J. N. Carter. Advances in automatic gait recognition. In *IEEE Face and Gesture Analysis 2004, FG04*, pages 11–16. IEEE CS Press, 2004. URL `http://eprints.soton.ac.uk/260069/`. Event Dates: 2004.

[28] L. O'callaghan, A. Meyerson, R. Motwani, N. Mishra, and S. Guha. Streaming-data algorithms for high-quality clustering. In *icde*, page 0685. IEEE, 2002.

[29] D. T. Pham, S. S. Dimov, and C. Nguyen. Selection of k in k-means clustering. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 219(1):103–119, 2005.

[30] H. Pohl and A. Hadjakos. Dance pattern recognition using dynamic time warping. *Sound and Music Computing*, 2010, 2010.

[31] P. T. Rushil Anirudh, Ahnaf Masroor. An online sampling for streaming video summarization. IEEE.

[32] S. V. Stevenage, M. S. Nixon, and K. Vince. Visual analysis of gait as a cue to identity. *Applied Cognitive Psychology*, 13(6):513–526, 1999. URL `http://eprints.soton.ac.uk/252496/`.

[33] K.-T. Tang, H. Leung, T. Komura, and H. P. Shum. Finding repetitive patterns in 3d human motion captured data. In *Proceedings of the 2nd international conference on Ubiquitous information management and communication*, pages 396–403. ACM, 2008.

[34] Q. Wang, P. Turaga, G. Coleman, and T. Ingalls. Somatech: an exploratory interface for altering movement habits. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, pages 1765–1770. ACM, 2014.

[35] J. Weiler. Mushi distortions, 2016. URL `http://jenweiart.com/mushi530/distortion.html`.