This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Expert-Driven Perceptual Features for Modeling Style and Affect in Human Motion

S. Ali Etemad, *Member, IEEE*, and Ali Arya, *Senior Member, IEEE*

***Abstract*—This paper presents a novel approach for modeling features of style and affect in human motion. Our approach is based on inputs collected from experienced animators. For this purpose, an interface is developed that allows for editing of motion sequences by adding a limited number of Gaussian radial basis functions (RBFs) to different joint trajectories in 3-D Cartesian space. Animators are asked to alter a neutral walking sequence to synthesize happy, sad, feminine, masculine, energetic, and tired variants. Through consolidating the sets of collected RBFs, we compute an expert-driven set of features that can transform neutral walks to the mentioned variations. Moreover, details regarding the use of posture versus movement features, the most frequently edited body joints, as well as shapes, intensities, and distributions of the edits are investigated and presented. Perception feedback from a group of nonexperts validates the proposed approach and the effectiveness, efficiency, scalability, and inversion of the proposed models. The perception study also sheds light on several aspects of perceiving style and affect from motion.**

***Index Terms*—Affect, energy, expert-knowledge, gender, human motion, inversion, movement, posture.**

I. Introduction

HUMAN motion and many variations with which it can be carried out have been widely studied. It has been broadly shown that viewers can accurately perceive many characteristics regarding the persons performing the actions. These perceivable variations include identity [1], gender [2], and affect [3], among others. For surveys on human motion analysis and perception, see [4] and [5].

Many studies have been carried out with the aim of modeling and describing stylistic/affective features in motion. Both computing and perception studies have been widely explored. The former, which are mostly aimed toward human–computer interaction and animation applications, have been based on a variety of methods such as machine learning [17], [32], component-based techniques [10], [11], [38], control techniques [22], [23], inverse kinematics modeling [30], and others. The latter are aimed at understanding the psychology and neuroscience of perceiving motion features and are based on subjective evaluation and comparative studies of recorded/generated motion data [35], [41]–[47]. Despite the accuracy of advanced computational approaches, they are often complex and computationally expensive. Furthermore, they require large training datasets to achieve acceptable generalization and are also sensitive to factors such as inclusion criteria. Comprehensive studies on perceptual outcomes and implications of such systems have also been mostly overlooked. On the other hand, although subjective observations and rule-based models are simple and intuitive, they are mainly difficult to mathematically model and apply to new data. Moreover, the relative perceptual significance and impacts of different observed features have not been widely explored. Finally, it has not been investigated whether subjective observations are scalable (in terms of intensity) or not.

This paper presents a simple, intuitive, and efficient approach for modeling stylistic and affective features in motion. We propose and utilize low-cost and controllable Gaussian radial basis functions (RBFs) as constructs of features in different categories of style and affect. Accordingly, we develop an interface that can be used to add multiple RBFs to motion sequences. Experienced animators are asked to use this interface and add up to three RBFs per degree of freedom (DOF) to a neutral walking sequence with the goal of generating stylistic sequences, namely, happy, sad, energetic, tired, feminine, and masculine walks. The neutral input sequence is synthesized through aligning and averaging multiple segmented neutral walks. This process ensures that the input is precisely neutral and free of personal walking styles. The aforementioned sets of RBF-based edits are recorded and combined to achieve the feature set for each category of style and affect. To evaluate the perceptual accuracy of the computed sets, they are applied to the neutral sequence and user studies are carried out. The study shows that the computed features are perceptually effective and accurate. The study also sheds light on how the mentioned motion variations are performed and perceived. Analysis of the edits reveals the shape, distribution, intensity, type, and other properties about the proposed method and computed features.

In Section II, the related works on the existing techniques for modeling motion features are studied. Section III presents the general model used for describing how features of style and affect are added onto actions, the data used in this research, and the RBFs used to model the features. In Section IV, the interface with which the features are recorded is described along with the method and process of recording and summarizing the features. Section V provides an analysis on the features and perception results, followed by a detailed discussion in Section VI.

Manuscript received August 24, 2015; revised December 1, 2015 and February 7, 2016; accepted February 21, 2016. This work was supported in part by the Natural Sciences and Engineering Council of Canada (NSERC) and the Ontario Centres of Excellence (OCE). Some of the data used in this paper were obtained from HDM05 dataset [34]. This paper was recommended by Associate Editor D. Wu.

The authors are with the School of Information Technology, Carleton University, Ottawa, ON K1S 5B6, Canada (e-mail: ali.etemad@carleton.ca; arya@carleton.ca).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org.
Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.
Digital Object Identifier 10.1109/THMS.2016.2537760

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

II. Related Work

On the topic of computational systems for modeling, synthesis, or control of motion features, various interesting and effective methods have been developed, mostly in the field of animation. Most such techniques use examples and datasets as an integral part of the solution. For example, in two of the earlier works in the field, Rose *et al.* [6] and Bruderlin and Williams [7] used relative editing of sequences to blend motion styles successive to time warping. In [18], Hsu *et al.* presented a method for human motion control using dynamic programming and segments of motion sequences.

Component-based methods have been utilized to decompose style features and actions. In [8], Liu *et al.* decomposed motion sequences into multiple subspaces, allowing animators to manipulate (tune, transfer, or merge) the style subspace. Urtasun *et al.* [9] developed an extended framework using principal component analysis (PCA) for real-time generation of motion sequences with increased realism. Kim and Neff [10] employed independent component analysis to decompose motion to submotion style components. Extracted style components were then applied to new motion clips.

Machine learning methods have been widely explored for modeling or translating style features. Hsu *et al.* [11] proposed a system that learned style translation models using linear time-invariant system identification following motion warping. Lee and Popović [12] proposed a system that used a limited number of examples to learn behavior styles and apply them to different scenarios. Torresani *et al.* [13] introduced a novel algorithm for generating stylistic motion sequences. Laban movement analysis was used in this research to perceptually depict the motion styles. Their model learned a regression problem from motion capture data and could subsequently interpolate/extrapolate styles. Lee *et al.* [14] proposed a new representation for motion data, called motion field. Their flexible approach used prerecorded datasets and could be used for real-time control of motion. In [15], Kwon and Shin presented an example-based online motion synthesis system successive to an analysis of the dataset that was used.

Samadani *et al.* [51] utilized probabilistic graphical models and dimensionality reduction along with different classifiers to model affective movements. Min *et al.* [16] proposed a model for synthesis and editing styles in human motion data. Their approach utilized preregistered samples from a database and could be used for style translation as well as synthesis of new styles. Probabilistic models proposed by Brand and Hertzmann [17] used a rich motion dataset to learn interpolated/extrapolated styles using a cross-entropy optimization framework for dance movements. Kingston and Egerstedt [49] also utilized optimization along with estimators to model and control expressed preferences in synthetic motions. Optimal control was used by LaViers and Egerstedt [50] in order to determine stylistic parameters for motion and analyze the quality. Last but not least, Pullen and Bregler [21] used keyframed data along with motion capture data to generate natural and "textured" motion. Their approach was based on intercorrelations of body motion.

Physics-based or control-based methods for modeling and control of motion have also been widely explored. Liu *et al.* [22] used dynamic simulations with motion capture data to model contact forces on characters based on user-indicated bounds. de Lasa *et al.* [23] developed motion controllers. Their method could model different types of actions and was robust towards changes in body parameters.

Interactive or user-driven editing of motion is a topic that closely relates to the study presented in this paper. Bruderlin and Calvert [24] developed a hybrid rule-based dynamic control system for gait generation. They also proposed a knowledge-driven set of procedures for synthesizing different styles of human running [25]. Amaya *et al.* [33] developed a model for making edits to movement for generating affect. Witkin and Popovic [26] proposed a framework in which animators could interactively apply edits in the form of deformations or warps to motion to create variations of input sequence. Neff and Kim [27] introduced a platform for stylistic editing of motion particularly focusing on wrists, ankles, center of mass, and pelvis. Arikan *et al.* [19] proposed a framework for synthesizing user-driven motion sequences based on a dataset of preannotated motion data. In [20], Arikan and Forsyth used a cut-and-paste approach for generating human motion data. Naturally, this approach was based on existing sequences. Challenges such as smooth transitions and natural-looking motion were addressed in this study.

As mentioned in Section I, we utilize Gaussian functions for the purpose of modeling style and effect. Gaussian processes have been previously used for motion processing. Wang *et al.* [28] employed Gaussian process latent variable model (GPLVM) for motion synthesis. Urtasun *et al.* [29] proposed locally linear GPLVM for learning stylistic motion and transitions. Grochow *et al.* [30] proposed and used scaled GPLVM for generating a variety of poses based on inverse kinematics.

The aforementioned computational techniques are valuable tools for modeling motion and affective/stylistic features. Nevertheless, despite their accuracy, they are not necessarily perceptually accurate or computationally efficient and often do not provide much insight into how the features impact the base motion and what their degrees of importance are. Rule-based and expert-driven systems, on the other hand, despite their intuitive nature and perceptual validity, are usually hard to mathematically model and quantify. Through the following sections, we present our method, which consolidates the advantages of both approaches.

III. Overview and Theory

A. Overview

Our approach for modeling and understanding style and affect in motion is based on a linear model describing the relationship between the action component of the motion and the stylistic features. This is described in the following subsection. The proposed method for modeling specific stylistic/affective features is based on having animators provide Gaussian RBFs that convert a neutral base motion into several stylistic variations. The RBF edits are collected and subsequently summarized to develop a feature set. The feature set is then applied to a neutral walking sequence and evaluated through a user perception study. A quantitative analysis of the feature set is also carried out, illustrating several informative traits with respect to each stylistic/affective variation and body part.

B. Linear Additive Model for Affective/Stylistic Motion

In human motion, the main action class is referred to as primary actions and the affects, styles, or attributes associated with the actor performing the actions are referred to as secondary themes [31], [32]. Accordingly, for a recorded sequence of motion, the set of spatiotemporal features that compose a primary theme is called primary features, and those responsible for generating the themes are called secondary features (SFs). Based on this definition, the following model has been proposed for describing the sets of primary and SF in motion [31, 32]:

$$\mathbf{Y} = \mathbf{P} + \sum_{i=1}^{r} \mathbf{w}_i.\mathbf{S}_i \tag{1}$$

where $\mathbf{Y}$ is the action as perceived, $\mathbf{P}$ and $\mathbf{S}$ are the primary and secondary themes (or features) respectively, r represents the number of secondary themes, and $\mathbf{w}$ is the weight associated with each secondary theme. In order to simplify the problem, it is often assumed that $r = 1$. In other words, combinational secondary themes such as young-tired or energetic-feminine are disregarded [32]. Therefore, from (1), we obtain $\mathbf{Y} = \mathbf{P} + w.\mathbf{S}$. The most accurate representation of primary and secondary themes is achieved through this model when the two themes are in the null-spaces of one another [48].

C. Motion Data

Based on the described theory, a stylistic motion sequence is composed of a neutral action and an added and scaled style component. Here, we use a neutral walking sequence as the input to which experts add affective/stylistic features. As a result, for the input, the primary theme $\mathbf{P}$ corresponds to walking and contains zero or very little amount of SF; in other words, $\mathbf{S} = 0$. The notion of neutral actions has been well explored before [7], [33].

To compute the neutral input in this research, we use motion capture data from the HDM05 dataset. This dataset contains motion data recorded using a marker-based Vicon motion capture system. The system utilizes 6–12 cameras to track 40–50 light-reflective markers located on a body suit. The cameras have very high temporal (up to 240 Hz) and spatial (less than 1 mm) resolutions. Five actors have performed multiple actions, among which are sequences of neutral walks. To the best of our knowledge, details regarding the age and gender of the actors are not disclosed. A detailed documentation for the dataset is available in [34].

Motion capture data can be represented by a number of consecutive multidimensional postures variable with time. Each posture is characterized with a finite number of markers corresponding to different regions or joints of the body to create a motion matrix. The motion matrix can either be characterized through the location of markers at each frame or instance of time, joint angles, or other means. We represent the motion matrix by

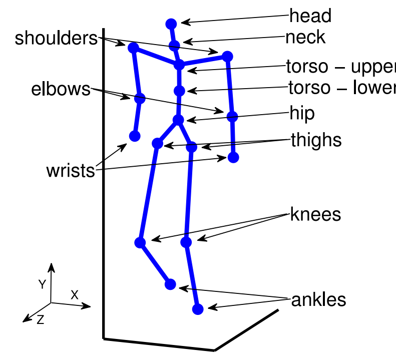

Fig. 1. 54 DOF model used for the human body in this study. This posture is from the neutral walk generated by warping and averaging multiple neutral walks from the HDM05 dataset.

$$\boldsymbol{\mathcal{D}} = [\boldsymbol{R}_1 \boldsymbol{R}_2 \ldots \boldsymbol{R}_n] \tag{2}$$

where $\boldsymbol{R}_j = [R_j^{(1)} R_j^{(2)} \ldots R_j^{(m)}]^T, R \in \mathbb{R}$. $\boldsymbol{R}_j$ represents the jth DOF of the motion sequence with a length of m frames or time instances, and n represents the number of DOFs in the motion matrix. The model used for the HDM05 data is composed of 96 DOFs, some of which belong to fingers and toes. We modified and removed these extra joints from the model, reducing it to 54 DOFs.

To record a neutral action sequence, actors are often asked to display minimal affective and stylistic behavior. Complete concealment of SFs, however, is not fully achievable. To maximize the neutrality of an action sequence and eliminate the personal walking styles of the actors, we propose averaging multiple neutral sequences. Accordingly, we manually segmented and extracted 16 two-step neutral walks from the HDM05 dataset. The length of these walk cycles were sufficient for correct perceptual studies [35]. We then aligned the segmented cycles using correlation optimized time warping (CoTW) [36], [37]. This time warping method has shown characteristics such as peak preservation and low distortion, making it superior to alternative methods commonly used for motion data. By averaging the segmented cycles, an average neutral walk was obtained. Fig. 1 illustrates a posture from this sequence where the 54 DOF model is observed. The sequence is 100 frames long. The accompanying video clip shows the neutral walk cycle. As seen in the video, there are no artefacts and the quality of the sequence is on par with the video of the average walker created in [38]. To further ensure the quality of this sequence, it is perceptually validated in Section V.

D. Radial Basis Functions as Constructs for Secondary Features

One approach to editing motion is to apply ad-hoc and freestyle transforms to existing motion trajectories. However, due to the highly varying nature of this approach, consolidation, analysis, and interpretation of such processes are very difficult if not impossible. To overcome this obstacle, we propose the use of standard mathematical functions, namely Gaussian RBFs as constructs for SFs. These functions are highly controllable and easy to analyze and study. Moreover, when presented in parametric format, Gaussian RBFs maintain low dimensionality. Such properties make RBFs perfect candidates for generation of SFs by animators. Moreover, it has previously been demonstrated that they are both numerically and perceptually suitable for modeling SFs [32].

A radial function $\phi\,(r)\,, \phi : \mathbb{R}^s \rightarrow \mathbb{R}$ is defined as a univariate function, where $r = \|t_2\|$, and $\|.\|_2$ is a norm operator such as the Euclidean norm. Accordingly, a Gaussian RBF is defined by

$$\varphi\left(t;\mu,\sigma^2\right) = \phi\left(\|t-\mu_2\|\right) = \exp\left\{\frac{-\|t-\mu\|_2^2}{2\sigma^2}\right\} \tag{3}$$

where μ is the mean, and σ^2 is the variance. We can model the SF trajectory of the ith DOF with a weighted sum of M RBFs, resulting in

$$\mathbf{\Phi}_i = \sum_{j=1}^{M} \alpha_j \varphi\left(t;\mu_j,\sigma_j^2\right). \tag{4}$$

where j denotes the RBF index, and α is the amplitude or intensity. Hence, the SF set for the *m*-frame long *n*-dimensional motion sequence is represented by $\{\mathbf{\Phi}_1, \mathbf{\Phi}_2, \ldots, \mathbf{\Phi}_n\}_{n\times m}^T$ or by the $n \times 3M$ parameter set matrix:

$$\mathbf{Q} = \begin{bmatrix} \{\alpha,\mu,\sigma^2\}_1^1 & \cdots & \{\alpha,\mu,\sigma^2\}_M^1 \\ \vdots & \ddots & \vdots \\ \{\alpha,\mu,\sigma^2\}_1^n & \cdots & \{\alpha,\mu,\sigma^2\}_M^n \end{bmatrix}. \tag{5}$$

In this model, the larger the M, the more accurate the modeled SF will be. However, there will be little added perceptual and numerical return after a certain point. It is, therefore, important to maintain an acceptable balance between complexity (number of RBFs) and accuracy. The variable is also dependant on the complexity of both action class and SF. A previous study [32] has shown that for the action class of walking and for SFs such as those studied in this work, *three* RBFs per DOF are sufficient for achieving a perceptually valid and numerically accurate model for SFs. We, therefore, utilize the notion and limit $M \leq 3$. Nonetheless, for more complex actions and secondary themes, larger values of M might be required.

IV. Interface and Data Collection

To facilitate the process of collecting the $\mathbf{Q}$ feature matrices from animators, a user interface capable of generating the RBFs and adding them to corresponding DOFs of motion was required. To the best of our knowledge, such a platform did not exist at the time that this study was being conducted. Subsequently, we developed an interface for this purpose. For simple integration and easier analysis of the results, the interface was developed in MATLAB using the *guide* functionality.

A snap shot of the graphical user interface (GUI) is presented in Fig. 2. Using the system, users can browse and load motion capture data in the form of *bvh* files (top-left section of the interface). A loaded motion capture file can then be animated (right side of the interface). For manipulating the motion, a particular skeleton joint and an accompanying Cartesian axis (corresponding to a DOF of the motion matrix) can be selected. The selected joint is highlighted in red as opposed to the other joints displayed in blue. For each DOF, up to three RBFs can be generated in Cartesian space. The three parameters (α, μ, σ^2) for each generated RBF are tuned using sliders and the corresponding RBF is interactively plotted (left side of the interface). Each set of generated RBFs can be summed to create a single curve. The sum of RBFs can then be applied to the animated motion sequence. When a set of RBFs has already been synthesized for a selected DOF, reselecting that DOF displays those RBFs for the user for possible further modification. Previously saved sets of RBFs can also be loaded and modified. Upon completion of a task, the parameter set ($\mathbf{Q}$) can be saved.

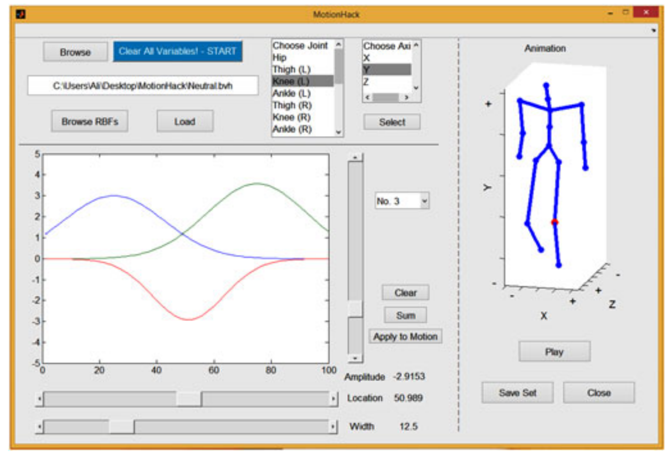

Fig. 2. The GUI developed and used to generate the features is depicted. The interface enables loading of motion capture files, generating RBFs, adding them to the sequence, and animating the original or modified sequence.

Twenty-seven subjects in two groups participated in this study. The first group was composed of 11 participants who were experienced with animation. They were either graduate students with related experience or employees of the private sector, working for related companies. Their mean age was 25.8 with a standard deviation of 4.2; nine were males and two were females. They were provided with paper-based description of the process and interface. Their task was to convert the average neutral walk (described earlier) to happy, sad, energetic, tired, feminine, and masculine styles using the interface. They were first asked to practice with the interface to ensure proper utilization of its different functionalities. They were then asked to convert the loaded neutral walk to each of the mentioned themes. They were instructed to use as many RBFs (up to three) per DOF, and modify as many DOFs as they felt required to complete the task. They could use the GUI to play the animation at any time in order to evaluate the synthesized and added features. Correcting or deleting of added features was permitted. Upon conversion of the neutral walk to each of the themes, the added features were saved and the interface was reset. The order in which the themes were generated by different users was randomized to prevent arrangement bias. The feature collection process was quite time consuming (between 90 and 120 min). Participants

TABLE I
VALIDATION OF THE GENERATED INPUT NEUTRAL WALK

Neutral	*Happy*	*Sad*	*Energetic*	*Tired*	*Feminine*	*Masculine*
75.0%	6.25%	6.25%	0.0%	6.25%	0.0%	6.25%

were compensated for their time. The second group of participants consisted of 16 individuals, 5 of whom were females and 11 were males. Their average age was 29.8 with a standard deviation of 10.9. They took part in providing perceptual feedback on the generated results for analyzing the different parameters involved in this study. They were naive toward human motion studies. A seven-point forced-choice paper-based questionnaire was used. The MATLAB-based interface as well as the animated sequences were presented on a desktop computer with 3 GB of RAM, a 2.8-GHz processor, and a 23.6-in 1080 HD LED screen. Ethics approval was secured from the Research Ethics Board of Carleton University for both components of the study.

V. RESULTS

Successive to data acquisition from the animators, based on (5), the feature set created by participant i for SF type j is denoted by $\mathbf{Q}_{i,j}$, where $i = 1 \dots 11$ and $j = 1 \dots 6$ ({*Happy*, *Sad*, *Energetic*, *Tired*, *Feminine*, *Masculine*}). The $\mathbf{Q}$ matrices are then averaged across the 11 animators (i). Since participants were given the choice of utilizing one, two, or three RBFs per up to 54 DOFs, some elements of $\mathbf{Q}_{i,j}$ can be zeros (some users chose not to modify a particular DOF or not to use all three RBFs at their disposal). As a result, averaging $\mathbf{Q}_{i,j}$ for a given j across all i will depend on the order of the three RBFs. To address this issue, we sort each $\mathbf{Q}_{i,j}$ matrix based on μ_l values in ascending order prior to the averaging step. In other words, the RBFs are sorted such that $\mu_1^k < \mu_2^k < \mu_3^k$, where k denotes the DOF. α_l^k and $\sigma^{2_l^k}$ are rearranged along with the associated μ_l^k values. This will ensure that the temporal order of the RBFs is maintained and averaging is meaningful. The output matrix is then used to analyze the features and create a single feature set for each secondary theme, which can be added to the neutral sequence to achieve a stylistic/affective sequence.

A. Validating the Input

As the first step toward analyzing the results, the second group of participants mentioned in Section IV validated the neutrality of the averaged input sequence. The goal is to ensure that the input sequence does not contain significant features belonging to any of the six secondary themes of interest, which could lead to skewing of the final results. Table I presents the perception outcome for the input. The sequence is mostly recognized as *neutral*, and the false classifications being distributed across different themes, points to the absence of a particular dominant secondary theme in the input.

B. Distribution

The animators were not asked to modify every DOF of the motion sequence. As a result, only the DOFs that they perceived to be more essential towards generation of the themes were modified. Fig. 3 illustrates the percentages of animators who modified each section of the body. We observe that except for sadness, in which all 11 animators modified the shoulders and neck, no other section of the body drew this kind of consensus for any theme. When generating sadness, the frequency for the head, upper torso, elbows, and wrists followed. For generating the happy, energetic, and tired themes, shoulders drew the most attention. Wrists, knees, neck, head, and elbows followed. For the energetic theme, elbows, wrists, knees, and head came next. In generating the tired theme, ankles, neck, head, and wrists came after the shoulders. For the feminine theme, hip, thighs, and ankles are modified the most, followed by the knees and lower torso. Finally, for the masculine theme, knees, ankles, shoulders and elbows drew the most attention. Among all six themes, no body part was left unmodified. We performed two-factor repeated measures analysis of variances (ANOVA) on the number of edits performed by the animators, with body parts and secondary themes as independent variables. The analysis showed significant effects for both body parts ($F(10, 100) = 9.83$, $p = 3 \times 10^{-10}$) and secondary themes ($F(5, 50) = 4.15$, $p = 0.003$), indicating that the features significantly differ with either factor. The analysis also shows significant interaction between the two factors ($F(50, 500) = 2.73$, $p = 2 \times 10^{-7}$).

We observe that for happy, sad, energetic, tired, and masculine the distribution of added RBFs per body part weighs in favor of upper-body regions, with sadness having the most upper-body to lower-body advantage. The only exception of the six categories is feminine in which the lower-body drew more attention compared to upper-body regions. A two-factor repeated measures ANOVA for upper versus lower body and theme shows significant effect for upper versus lower body ($F(1, 10) = 21.78$, $p = 0.0009$) as well as theme ($F(5, 50) = 4.15$, $p = 0.003$), with significant interaction ($F(5, 50) = 4.62$, $p = 0.002$).

In terms of left/right half of the body, animators chose to treat shoulders, elbows, wrists, thighs, knees, and ankles (which are composed of left and right sides) symmetrically, meaning in all cases where these joints were altered, both left and right joints were modified. A two-factor repeated measures ANOVA for left versus right sides of the body and theme shows no significant effect for left/right ($F(1, 10) = 0.50$, $p = 0.50$), indicating the two sides were not altered differently. The effect of theme was significant ($F(5, 50) = 5.71$, $p = 0.0003$), and no interaction was observed ($F(5, 50) = 0.50$, $p = 0.78$).

C. Feature Properties

1) Common Shapes: The sets of RBFs created and used by the animators for modifying different DOFs of the motion can be categorized into four major types or shapes. Fig. 4 illustrates these four commonly used features. Features types 1 and 2 are composed of one single RBF, while feature types 3 and 4 are composed of two. The blue and green curves show the individual RBFs while the red represents the sum. While shapes other than these were also observed in the results, such features were nonrecurring.

In general, feature type 1 is utilized to add a spatial offset to a motion trajectory of a particular DOF. These features often

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

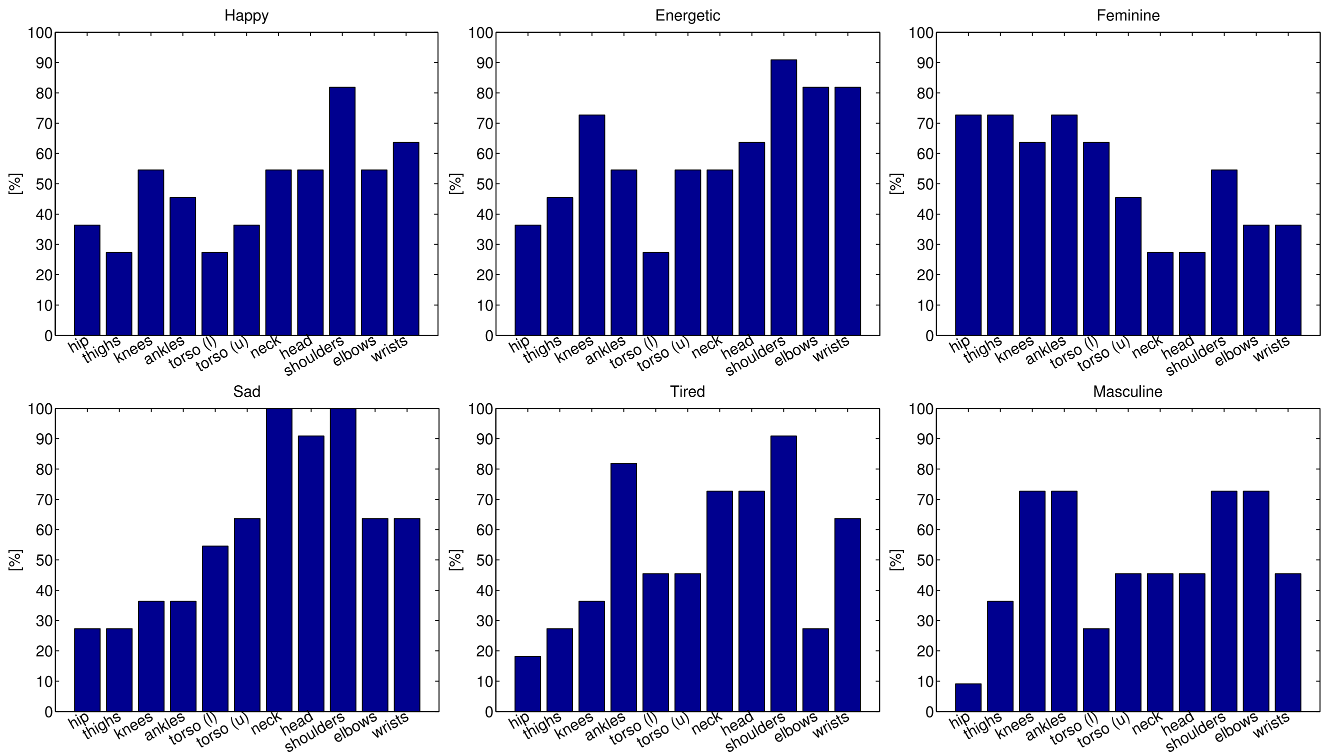

Fig. 3. The histograms show the percentage of animators that have modified each part of the body.

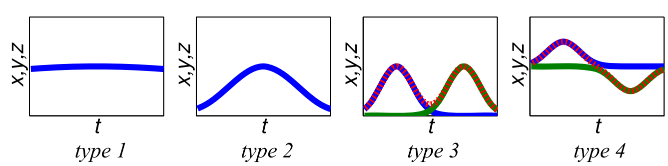

Fig. 4. The most common shapes throughout the feature set are presented.

TABLE II
PERCENTAGE OF MOVEMENT VERSUS POSTURE FEATURES USED BY ANIMATORS TO GENERATE DIFFERENT THEMES

	Happy	*Sad*	*Energetic*	*Tired*	*Feminine*	*Masculine*
% Movement	71.4	24.5	86.0	22.9	59.0	67.3
% Posture	28.6	75.5	14.0	87.1	41.0	32.7

modify the body posture and not motion trajectories. An applied example of this is when the head is tilted down. Feature type 2 is utilized when the animator intends to change the movement of a particular joint by increasing and subsequently decreasing it during a certain time period, or vice versa. In other words, this feature can be used to add local maxima/minima where none exists, or to increase/decrease the amplitude of existing minima/maxima. For example, this feature can be used to increase the movement of the right elbow. Similarly, feature type 3 is used to repeat this process at two different points. For example, in a two-step walk, like the one being used as the input in this study, feature type 3 can be used to add two forward sways to the head, one with each step. Finally, feature type 4 is utilized to add opposing extrema to a trajectory. For example, it can be used to add sway to the hip joint where with the right step, the hip shows increased movement toward the left, and with the left step, it sways toward the right. For mirroring of the features for left/right joints, vertically mirrored ($-\alpha$) versions of the four types of features are used.

2) Posture Versus Movement: The use of feature type 1 versus the other three types requires further analysis. In biological motion studies, style and affect are reported to be composed of two general types of spatiotemporal features, namely posture and movement (dynamics) [39], [40]. Additionally, speed or uniform time need to be considered, which we will discuss in Section VI. Posture features are those that often remain unchanged throughout the sequence. In effect, they alter the initial posture of the body with which the motion is carried out. Movement features are changes to the motion trajectories that vary throughout the sequence. Table II presents the percentage of DOFs of the neutral walk that have been modified using posture features (type 1) versus movement features (type 2, type 3, and type 4). We observe that sad and tired sequences, which are in fact quite similar to each other, are mostly generated with posture features while the rest are mostly created by movement features. The maximum relative percentage of movement features is utilized for energetic while the maximum relative percentage of posture features is used for tired walk.

3) Features: Table III presents the ten most frequently generated features for each theme. As discussed, two general feature types, posture and movement, are used, which we refer to in the table as *tilted along* and *increased/decreased swing along*

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE III
TOP TEN FREQUENTLY USED FEATURES FOR GENERATION OF THE THEMES

		Themes					
		Happy	*Sad*	*Energetic*	*Tired*	*Feminine*	*Masculine*
Common features	*1*	Shoulders: increased swing along *Z*	Shoulders: tilted along –*Y*	Knees: increased swing along *Y*	Shoulders: tilted along –*Y*	Hip: increased swing along *X*	Shoulders: increased swing along *Z*
	2	Wrists: increased swing along *Z*	Head: tilted along –*Y*	Head: tilted along +*Y*	Head: tilted along –*Y*	Ankles: tilted, R along +*X*, L along –*X*	Knees: tilted, R along –*X*, L along +*X*
	3	Knees: increased swing along *Y*	Neck: tilted along –*Y*	Elbows: increased swing along *X*	Ankles: decreased swing along *Z*	Knees: tilted, R along +*X*, L along –*X*	Ankles: tilted, R along –*X*, L along +*X*
	4	Head: increased swing along *X*	Shoulders: decreased swing along *Z*	Elbows: increased swing along *Z*	Head: tilted along +*Z*	Torso L: increased swing along *X*	Elbows: tilted, R along –*X*, L along +*X*
	5	Wrists: increased swing along *X*	Wrists: decreased swing along *Y*	Shoulders: increased swing along *Y*	Wrists: tilted along –*Y*	Thighs: tilted, R along +*X*, L along –*X*	Ankles: increased swing along *Y*
	6	Hip: increased swing along *X*	Torso U: tilted along –*Y*	Shoulders: increased swing along *Z*	Ankles: decreased swing along *Y*	Torso U: increased swing along *X*	Knees: increased swing along *Y*
	7	Knees: increased swing along *X*	Head: tilted along +*Z*	Wrists: increased swing along *Z*	Neck: tilted along –*Y*	Thighs: increased swing along *Y*	Head: increased swing along *Z*
	8	Neck: tilted along +*Y*	Neck: tilted along +*Z*	Wrists: increased swing along *X*	Neck: tilted along +*Z*	Shoulders: tilted along –*Y*	Neck: increased swing along *Z*
	9	Head: tilted along +*Y*	Elbows: tilted along –*Y*	Thighs: increased swing along *Z*	Torso L: tilted along +*Z*	Elbows: increased swing along *X*	Shoulders: increased swing along *Z*
	10	Head: tilted along –*Z*	Hip: tilted along –*Y*	Thighs: increased swing along *Y*	Torso U: tilted along +*Z*	Wrists: increased swing along *X*	Elbows: increased swing along *Z*

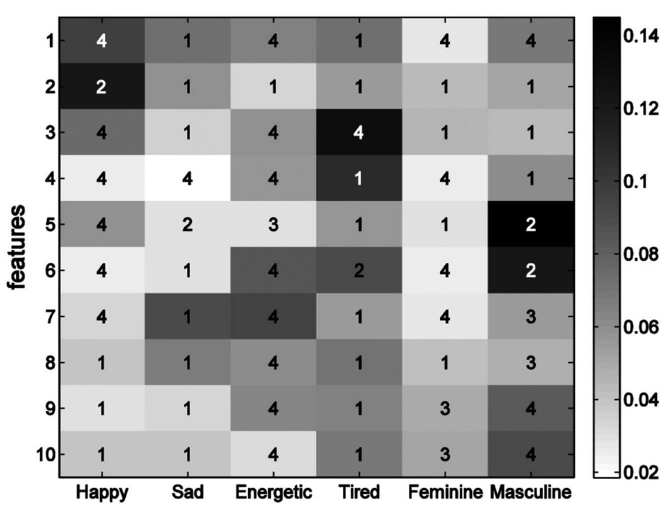

Fig. 5. Normalized intensities of features from Table III. The numbers in the cells represent the shape of feature based on Fig. 4 (some numbers appear in white for readability purposes given their dark backgrounds).

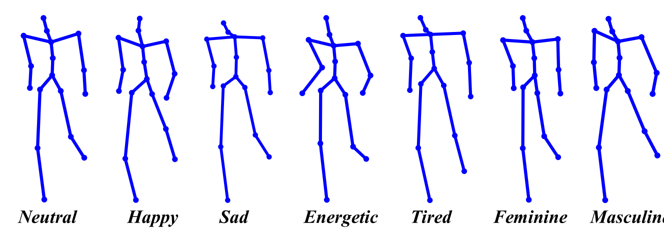

Fig. 6. Frames from the generated affective/stylistic sequences.

respective axis. *R* and *L* denote left or right joints. Utilizing all of the features mentioned in Table III and other less frequently used ones (that are not mentioned in this table) results in successful generation of the intended themes. However, we suggest that generating the themes can be achieved using only a few of these features.

Fig. 5 illustrates the intensities (amplitudes) of the features mentioned in Table III, normalized by the height of the averaged input walker. Normalization is carried out since the height of the walker can affect the magnitude of added features. For example, increased arm swing for a tall walker is greater than that of a short one. The figure shows that the most frequent features do not necessarily have the highest intensities. We observe that increased wrist swing along *Z* in happy, decreased ankle swing along *Z* in tired, and increased ankles swing along *Y* in masculine have the greatest intensities. This observation is logical since movement features are spatiotemporally (but not necessarily perceptually) dominant compared to posture features. As a result, modifying them would require stronger RBFs. Moreover, the three mentioned features belong to wrists and ankles, which generally show the most spatiotemporal movement in walking [41]. Furthermore, the numbers in Fig. 5 present the shapes of features according to Fig. 4.

The accompanying video clip shows the input and the affective/stylistic walks generated using the 10 features. Fig. 6 illustrates a frame from the neutral input walk and the corresponding frames after addition of the RBF-based features. Subjectively, high-quality animation is achieved.

D. Perception and Scalability

As mentioned in Section IV, to evaluate the performance and parameters of the proposed method, a second group of subjects participated in this study. The same group provided the information required for validating the input neutral sequence (see Section V-A). They were asked to rate the amount of affect or style in the displayed sequences generated using the computed feature sets. Given the premise that the animators produced the secondary themes in question, a forced-choice

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

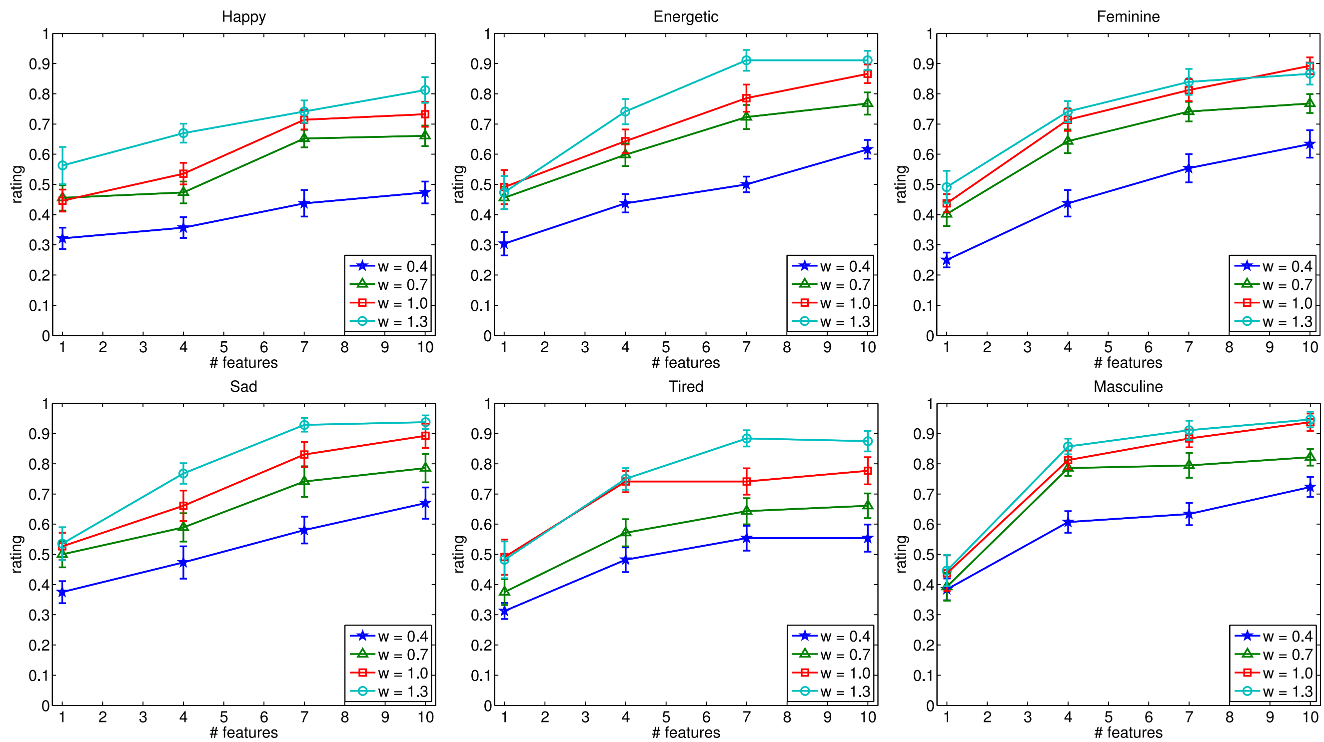

Fig. 7. Perception ratings for the two variables: weight and number of features. Error bars represent standard errors.

TABLE IV
RESULTS FOR TWO-WAY REPEATED MEASURES ANOVA WITH WEIGHT AND NUMBER OF FEATURES AS THE TWO INDEPENDENT VARIABLES

		Factor					
		Weight		Number of features		Interaction	
		F(3,45)	*p*	*F*(3,45)	*P*	F(9,135)	*p*
Themes	*Happy*	24.60	<0.0001	39.49	<0.0001	1.54	n.s.
	Sad	63.89	<0.0001	18.46	<0.0001	2.32	<0.05
	Energetic	88.62	<0.0001	23.84	<0.0001	2.17	<0.05
	Tired	33.39	<0.0001	21.22	<0.0001	1.61	n.s.
	Feminine	79.52	<0.0001	17.53	<0.0001	0.79	n.s.
	Masculine	102.56	<0.0001	14.11	<0.0001	2.70	<0.001

approach was use. The sequences were generated when one, four, seven, and ten features were used. The weight parameter (w) from (1) was another variable in this experiment with the values of $w = 0.4, 0.7, 1.0, 1.3$. Fig. 7 illustrates the results. Error bars represent standard errors ($\mathrm{SD}/\sqrt{n}$). Generally, a direct relation between the two factors and the perceptual ratings is observed. We performed two-factor repeated measures ANOVA for weight and number of features. The outcome is presented in Table IV, where for very theme, both weight and number of features show significant effect at the $p < 0.0001$ level. There is no significant interaction between the two factors for happy, tired, and feminine themes, while sad, energetic, and masculine showed significant interaction at $p < 0.05$, $p < 0.05$, and $p < 0.001$ levels, respectively. This analysis shows that the features are scalable and can be used to develop different levels of style and affect. Moreover, it indicates that while utilizing a subset of the features can produce the desired secondary theme to some extent, larger subsets result in more accurate representations.

E. Inversion

To further investigate the proposed method and the computed set of features, we tested negative w values. Surprisingly, we noticed the appearance of features belonging to the opposite themes. For example, when a negative weight was applied to happy features, indicators of sad motion were observed. It should be noted that the negative weight parameter only inverts α and not the other parameters.

We further tested this concept by applying a weight of $w = -0.7$ to all the ten features and collected perceptual ratings. Similar to validating the results with positive weights, given the premise of scalability and the notion of inversion in the literature [35], a forced-choice approach was used. The audience were asked to select a rating between -7 to $+7$, with $+7$ denoting the maximum original theme, 0 denoting neutral, and -7 denoting the maximum opposite theme. The results are presented in Fig. 8 where the audience perceived the opposite theme in all cases. Error bars represent standard errors. For some themes, this inversion effect seemed to be stronger. This inversion effect was especially stronger in masculine, happy, and feminine compared to energetic and tired themes.

VI. DISCUSSION

A. Features and Perception

The main goal of this research was to develop a set of features that model the affective and stylistic variations with little

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

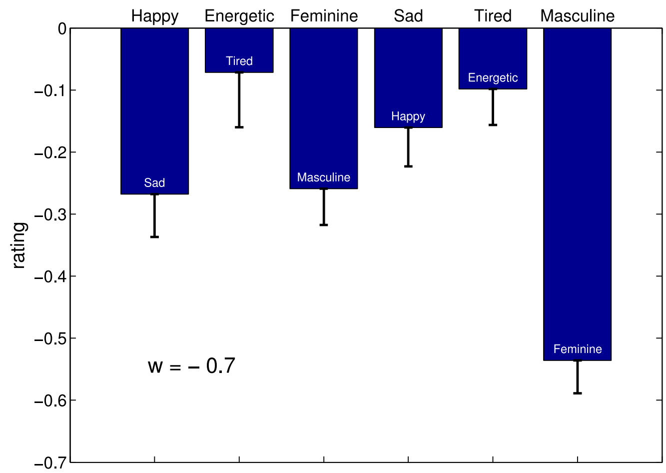

Fig. 8. Perceptual ratings for negative weights show theme inversion. Error bars represent standard errors.

computational effort and from a perceptual standpoint. We validated the results using perception feedback from inexperienced participants (see Section V-D). We have shown that only a subset of the features can induce the desired perception for the audience, which, compared to existing style translation or synthesis techniques that often use complex machine learning methods, is significantly low cost in terms of computational complexity. Moreover, the features are scalable and can be tuned to increase/decrease the induced perception of the themes. Our study also describes the nature of the features in terms of significance, distribution across the body, type, and shape.

There have been many studies in the literature attempting to describe features responsible for presentation and perception of affects/styles in motion. For example, it was observed in [41] that the following features are present in happiness: inclination of shoulders, lateral movements of hands/arms, arms being stretched out frontal, arms being crossed in front of chest, opening/closing of hands, and sideway positioning of back of the hands, while for sadness: collapse of the upper body, downward bending of the head, sideways bending of the head, lateral movements of hands/arms, and sideway positioning of back of the hands were reported. The study also suggests that in happy motion, more movement features are present compared to posture features. These features do not contradict our findings and our proposed feature sets. Identical features in both happy and sad themes reported in the mentioned study, seem unusual, and the number of reported features are low, as features related to legs and feet are ignored. Nevertheless, there are reports that suggest that emotions are mostly conveyed by the upper body [39]. Our results point to a similar dominance of upper body features for happiness and sadness.

In [43], it was illustrated that head inclination (posture) is central for sadness and that, otherwise, movement features are dominant for emotions. Table III in this paper shows a similar trend.

In [44], regression analysis and PCA were used to analyze emotions in recorded sequences. It was reported that posture features are more dominant in sadness compared to happiness. In terms of movement, for happiness, shoulders showed the most change followed by elbows, while for sadness, the same order was observed followed by the hips. Using PCA, for happiness, elbows displayed the only significant change, while for sadness the knees were dominant followed by hips.

For energy-related features, we were unable to find suitable studies that investigate motion features. For gender-related features, however, Troje [38] illustrated that movement-related features are generally more critical compared to posture-related ones. Increased hip sway in feminine and head sway in masculine walks have been suggested in [45]. It has been illustrated that body sway and lateral movements are dominant compared to posture changes and that increased sway of hip for feminine and shoulders for masculine are dominant features [46]. Our method reconfirms and expands on these observations.

In regard to the impact of the number and weight of features (see Section V-D), as the number of features increases, the perceptual ratings increase as well. In most cases, the increase in the ratings is more evident in the first few number of features and approaches a steady state as the number of features grows. This steady state corresponds to the maximum achievable average rating, in most cases approximately 0.9. The ratings are a bit lower for the happy theme (maximum of approximately 0.8). The reason for the existence of a steady state, we believe, is the fact that the features are arranged in the order of frequency, i.e. perceptual significance. This means that higher order features are less important and convey less information regarding the themes, resulting in insignificant changes to perception ratings. Similarly, the ratings increased with the increase of w. In most cases, the difference between $w = 0.4$ and $w = 0.7$ is greater than the difference between other consecutive pairs of weights, for example, $w = 10$ and $w = 1.3$. This property can be associated with the ratings for larger values of w approaching the maximum possible, i.e., 1.0. The 16 participants, however, indicated a decrease in visual quality for $w = 1.3$ and, in some cases, even for $w = 1.0$. This is due to the fact that the animators sometimes exaggerated in generation of the features. As a result, we opt for weights of approximately two-third for animation purposes. For example, $w = 0.7$ was used in Fig. 6 and the accompanying video clip. In addition, for analysis of the inversion effect, the same weight was utilized.

Regarding the loss of visual quality for larger values of w, an interesting observation was that as w approached values greater than 1, the audience stressed more loss of visual quality for happy, energetic, feminine, and masculine themes compared with sad and tired. The reason for this can be the fact that most features used for the former themes are movement-based. Exaggerating movement features results in unbound and unnatural motion, which can become distracting and unappealing. Posture features, on the other hand, when exaggerated, do not seem as unappealing as movement features since they only bring about changes in the structure.

For a broader analysis of posture versus movements features, we draw on the notion of activation and pleasantness in motion. We refer the reader to Russell's model or circumplex for affect, which uses activation and pleasantness as its two axes [42]. This model only discusses affects and not energy or gender.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

However, we can assume that, despite energetic actions often being associated with pleasant moods, an exclusively energetic action does not necessarily vary on the pleasantness axis while being positive on the activation axis. The same can be said about a tired action where it is negative on the activation axis and does not convey any information in terms of pleasantness given a happy-tired person can easily be realized. Similarly, for feminine and masculine, pleasantness does not play a role, and a particular judgment on activation is hard to make. However, due to exaggerated features, as well as being associated with faster gait cycles [35], slightly positive activation can be assumed for gender-related themes. In Table II, we observed that happy, energetic, feminine, and masculine variations are dominant in the use of movement features, while sad and tired variations make more use of posture features. Accordingly, we can conclude that, generally, themes with high activation are mostly generated using movement features whereas in themes that are low on activation, posture has a more important role. From a different standpoint, we can argue that *activity* requires *movement*. It is, therefore, logical for themes associated with higher activity to be associated with and represented by movement features. Where motion appears inactive, on the other hand, it is logical for the alternative type of feature, i.e., posture, to take the dominant role.

B. Time

All of the synthesized affective/stylistic actions had the same temporal length as the neutral input. In other words, the timing of the output sequences were not controllable through the GUI. Generally, there are two types of time features: uniform and nonuniform (also referred to as nonlinear). It has been previously shown that nonuniform temporal features are manifested as changes in movement [40]. Let us assume that a motion trajectory (or sequence) is altered such that its first half is linearly compressed by x frames and its second half is linearly stretched by the same number of frames (x). While the overall length of the resulting trajectory is preserved, there will be nonlinear temporal modifications applied to the trajectory. These nonlinear modifications can be interpreted as a spatiotemporal curve, i.e., movement feature. This type of time feature has been taken into account through the process of this research since movement changes in motion have been collected and analyzed. A uniform or linear time feature, on the other hand, is available when the entire trajectory is linearly stretched or compressed to achieve a new length. This type of time feature is simple to calculate or even estimate. Stylistic and affective variations of motion that are positive on the activation axis of Russell's model of affect [42] are often faster (shorter) compared to the neutral version, while lower activation motions are slower (longer). Moreover, it has been previously documented that viewers can easily recognize emotions of speed-matched affective sequences [44], [47] as posture and movement features alone provide sufficient cues for perception. It is, therefore, safe to conclude that the set of features derived in this study, or a subset of them, can successfully be employed to synthesize scalable affective/stylistic features, and speed alterations can subsequently be applied to the derived sequences [40] using linear operations such as uniform time warping [37].

C. Inversion

The illustrated inversion effect can have significant implications for psychophysics and multimedia applications. Similar effects have been previously addressed in the literature. For example, Barclay *et al.* [35] illustrated that a feminine walker is perceived as male, and *vice versa*, when the stimuli is inverted. Our approach slightly differs, however, since: 1) the affect/style features alone are inverted rather than the entire sequence (along with the walker); 2) features of each joint are inverted along their local axis and not the global axis. While deriving a detailed and accurate neural or psychological model that can describe this phenomenon requires in-depth study of the brain and its functionality, we can speculate that the existence of *some* inverted features in opposing themes can partially explain this phenomenon (see Table III). This would especially seem sensible for opposing posture features, rather than movement features. For example, in energetic and tired themes, we have tilted head along $+Y$ and tilted head along $-Y$, respectively. These features will convert to one another should they be spatially inverted. For increased/decreased swing (movement) features, further investigation is required.

D. Generalization

One of the benefits of the studied set of features is that it uses notions of increased/decreased swing and tilt, which are descriptions derived from the mathematical RBF-based feature set. The use of only two general and time-independent features increases the generalization of the models to actions other than walking. Nevertheless, the precise values of the parameters (α, μ, σ^2) will most likely need to be computed accordingly.

E. Keyframing

While our approach bears some similarities to keyframing, there are certain advantages to our approach. By restricting the animators to utilizing a limited number of controlled edits in the form of RBFs, features can be consolidated. As a result, conclusions regarding the nature of the features as a whole can be drawn that help with better understanding as well as generalization of the feature set. Moreover, our approach results in scalable features, which would not necessarily apply to a keyframing approach.

F. Limitations and Future Work

It should be pointed out that our method is founded on a perceptual approach and is not motion-data-driven. This means that our proposed feature set is not necessarily the same set of features that a feature extraction method would discover. Traits that are observed in generic stylistic motions are often the result of complex biological or evolutionary behaviors influenced by factors such as individual characteristics, culture, context, and more [4]. However, our study presents a set of features that, when applied to walking motion, can synthesize and convey the desired secondary themes based on what animators perceive to be crucial features. Further investigation of the correlations or differences between our proposed feature sets and those extracted from motion datasets can shed some light on the biological significance of the presented work.

As the proposed features are in the form of spatiotemporal transforms in Cartesian space, a limitation of our approach is that bone lengths are not necessarily preserved after the edits are applied. Nevertheless, the animated output sequences did not show any significant violations of bone lengths. This is perhaps because animators have taken the necessary measures to use edits that do not cause bone length variations, thus synthesizing naturally appearing sequences. Similarly, should different skeleton models or builds be used, adjustments might be required to achieve natural and perceptually valid motion. However, the underlying concepts of the features can be generalized and extended to other skeleton models (for example increased shoulder swing along Z). The proposed method only utilized walking as the action class. Therefore, while the general concepts of the proposed features would most likely be applicable to other action classes as discussed in the previous subsection, further investigation is required to determine the modifications required for the approach to apply to other action classes.

Another limitation of our proposed work is that if the input motion sequence contains a pre-existing secondary theme, the proposed feature set may not necessarily generate the desired output style. We believe further investigation is required to explore whether or not the proposed set of features can successfully synthesize the respective secondary themes with pre-existing secondary themes.

As described earlier, forced-choice questions were used in the validation step of the study, which might have influenced (most likely reduced) the divergence of the perception results. Alternatively, open questions or multiple forced-choice scales per question could be used, which would determine whether other themes are perceived from the produced sequences or not, and if so, to what extent.

For future work, we will investigate the use of the proposed features on different skeleton models and different actions to evaluate the methodological feasibility and perceptual significance of the results. The notion of pre-existing secondary themes will be investigated as well. Finally, in addition to analyzing this work with respect to a data-driven approach, we will utilize the proposed feature set to develop an automated classifier for recognition and retrieval of motion sequences. Comparing such a classification system with systems developed using motion datasets can shed light on the basis of the proposed perceptual features, as well as their generalization with respect to recorded motion sequences.

VII. Conclusion and Summary

We have proposed and implemented a method for recording expert-driven features for synthesis of happy, sad, energetic, tired, feminine, and masculine variations in motion. A user interface was developed using which edits in the form of Gaussian RBFs could be synthesized and applied to motion. A group of experienced animators converted a neutral walk into the mentioned variations. The features were recorded and compiled to derive a feature set. The advantages of our proposed method are as follows.

1) There have been many studies on the subject of modeling affect/style in motion, which are mostly based on recorded sequences and machine intelligence techniques. Our method, on the other hand, is based on opinions of experienced animators, making the proposed models efficient while very effective.
2) To the best of our knowledge, ranking or relative significance for reported features in affective/stylistic motion is not available in the literature, especially to the depth presented in this paper. This property allows for the most important features to be utilized for effective and efficient synthesis of affective/stylistic features.
3) Our proposed system presents mathematical and tunable basis for creating affect/style features and the model allows for scaling of the reported features. Furthermore, utilizing a subset of the proposed set of features leads to high-perception ratings.
4) Our approach is simple to expand to other classes of affect/style as well as actions.
5) Analysis of the findings adds to the existing body of knowledge on execution and perception of affect and style in motion. Details regarding body-wise distribution of features, types, intensities, frequencies, and shapes of features, are presented and discussed.
6) Our model explores and partially describes the inversion effect. However, we believe further studies are required in this regard.

References

[1] J. E. Cutting and L. T. Kozlowski, “Recognizing friends by their walk: Gait perception without familiarity cues,” *Bull. Psychonomic Soc.*, vol. 9, no. 5, pp. 353–356, 1977.

[2] L. T. Kozlowski and J. E. Cutting, “Recognizing the sex of a walker from a dynamic point light display,” *Perception Psychophys.*, vol. 21, no. 6, pp. 575–580, 1977.

[3] W. H. Dittrich, T. Troscianko, S. Lea, and D. Morgan, “Perception of emotion from dynamic point-light displays represented in dance,” *Perception*, vol. 25, no. 6, pp. 727–738, 1996.

[4] R. Blake and M. Shiffrar, “Perception of human motion,” *Annu. Rev. Psychol.*, vol. 58, pp. 47–73, 2007.

[5] A. Kleinsmith and N. Bianchi-Berthouze, “Affective body expression perception and recognition: A survey,” *IEEE Trans. Affective Comput.*, vol. 4, no. 1, pp. 15–33, Jan.–Mar. 2013.

[6] C. Rose, M. F. Cohen, and B. Bodenheimer, “Verbs and adverbs: Multidimensional motion interpolation,” *IEEE Comput. Graphics Appl.*, vol. 18, no. 5, pp. 32–40, Sep./Oct. 1998.

[7] A. Bruderlin and L. Williams, “Motion signal processing,” in *Proc. ACM SIGGRAPH*, 1995, pp. 97–104.

[8] G. Liu, Z. Pan, and Z. Lin, “Style subspaces for character animation,” *Comput. Animation Virtual Worlds*, vol. 19, nos. 3/4, pp. 199–209, 2008.

[9] R. Urtasun, P. Glardon, R. Boulic, D. Thalmann, and P. Fua, “Style-based motion synthesis,” *Comput. Graphics Forum*, vol. 23, no. 4, pp. 799–812, 2004.

[10] Y. Kim and M. Neff, “Component-based locomotion composition,” in *Proc. ACM SIGGRAPH/Eurographics Symp. Comput. Animation*, 2012, pp. 165–173.

[11] E. Hsu, K. Pulli, and J. Popović, “Style translation for human motion,” *ACM Trans.. Graphics*, vol. 24, no. 3, pp. 1082–1089, 2005.

[12] S. J. Lee and Z. Popović, “Learning behavior styles with inverse reinforcement learning,” *ACM Trans. Graphics*, vol. 29, no. 4, p. 122, 2010.

[13] L. Torresani, H. Peggy, and C. Bregler, "Learning motion style synthesis from perceptual observations," in *Proc. Int. Conf. Adv. Neural Inf. Process. Syst.*, 2006, pp. 1393–1400.

[14] Y. Lee, K. Wampler, G. Bernstein, J. Popović, and Z. Popović, "Motion fields for interactive character locomotion," *ACM Trans. Graphics*, vol. 29, no. 6, p. 138, 2010.

[15] T. Kwon and S. Y. Shin, "Motion modeling for on-line locomotion synthesis," *Proc. ACM SIGGRAPH/Eurographics symp. Comput. Animation*, 2005, pp. 29–38.

[16] J. Min, H. Liu, and J. Chai, "Synthesis and editing of personalized stylistic human motion," in *Proc. ACM SIGGRAPH Symp. Interactive 3D Graphics Games*, 2010, pp. 39–46.

[17] M. Brand and A. Hertzmann, "Style machines," in *Proc. 27th Annu. Conf. Comput. Graphics Interactive Tech.*, 2000, pp. 183–192.

[18] E. Hsu, S. Gentry, and J. Popović, "Example-based control of human motion," in *Proc. ACM SIGGRAPH/Eurographics Symp. Comput. Animation*, 2004, pp. 69–77.

[19] O. Arikan, D. A. Forsyth, and J. F. O'Brien, "Motion synthesis from annotations," *ACM Trans. Graphics*, vol. 22, no. 3, pp. 402–408, 2003.

[20] O. Arikan and D. A. Forsyth, "Interactive motion generation from examples," *ACM Trans. Graphics*, vol. 21, no. 3, pp. 483–490, 2002.

[21] K. Pullen and C. Bregler, "Motion capture assisted animation: Texturing and synthesis," *ACM Trans. Graphics*, vol. 21, no. 3, pp. 501–508, 2002.

[22] L. Liu, K. Yin, M. van de Panne, T. Shao, and W. Xu, "Sampling-based contact-rich motion control," *ACM Trans. Graphics*, vol. 29, no. 4, p. 128, 2010.

[23] M. de Lasa, I. Mordatch, and A. Hertzmann, "Feature-based locomotion controllers," *ACM Trans. Graphics*, vol. 29, no. 4, p. 131, 2010.

[24] A. Bruderlin and T. W. Calvert, "Goal-directed, dynamic animation of human walking," in *Proc. ACM SIGGRAPH*, 1989, pp. 233–242.

[25] A. Bruderlin and T. Calvert, "Knowledge-driven, interactive animation of human running," in *Proc. Conf. Graphics Interface*, 1996, pp. 213–221.

[26] A. Witkin and Z. Popovic, "Motion warping," in *Proc. ACM SIGGRAPH*, 1995, pp. 105–108.

[27] M. Neff and Y. Kim, "Interactive editing of motion style using drives and correlations," in *Proc. ACM SIGGRAPH/Eurographics Symp. Comput. Animation*, 2009, pp. 103–112.

[28] J. M. Wang, D. J. Fleet, and A. Hertzmann, "Multifactor Gaussian process models for style-content separation," in *Proc. 24th Int. Conf. Mach. learning*, 2007, pp. 975–982.

[29] R. Urtasun, D. J. Fleet, and N. D. Lawrence, "Modeling human locomotion with topologically constrained latent variable models," in *Proc. 2nd Conf. Human Motion–Understanding, Model., Capture Animation*, 2007, pp. 104–118.

[30] K. Grochow, S. L. Martin, A. Hertzmann, and Z. Popović, "Style-based inverse kinematics," *ACM Trans. Graphics*, vol. 23, no. 3, pp. 522–531, 2004.

[31] S. A. Etemad and A. Arya, "Modeling and transformation of 3d human motion," in *Proc. 5th Int. Conf. Comput. Graphics Theory Appl.*, 2010, pp. 307–315.

[32] S. A. Etemad and A. Arya, "Classification and translation of style and affect in human motion using RBF neural networks," *Neurocomputing*, vol. 129, pp. 585–595, 2014.

[33] K. Amaya, A. Bruderlin, and T. Calvert, "Emotion from motion," in *Proc. Conf. Graphics Interface*, 1996, pp. 222–229.

[34] M. Müller, T. Röder, M. Clausen, B. Eberhardt, B. Krüger, and A. Weber, "Documentation mocap database HDM05," Univ. Bonn, Bonn, Germany, Tech. Rep. CG-2007-2, 2007.

[35] C. D. Barclay, J. E. Cutting, and L. T. Kozlowski, "Temporal and spatial factors in gait perception that influence gender recognition," *Perception Psychophys.*, vol. 23, no. 2, pp. 145–152, 1978.

[36] N. P. Vest Nielsen, J. M. Carstensen, and J. Smedsgaard, "Aligning of single and multiple wavelength chromatographic profiles for chemometric data analysis using correlation optimised warping," *J. Chromatography A*, vol. 805, no. 1, pp. 17–35, 1998.

[37] S. A. Etemad and A. Arya, "Correlation-optimized time warping for motion," *Vis. Comput.*, vol. 31, pp. 1569–1586, 2015.

[38] N. F. Troje, "Decomposing biological motion: A framework for analysis and synthesis of human gait patterns," *J. Vision*, vol. 2, no. 5, pp. 371–387, 2002.

[39] A. Normoyle, F. Liu, M. Kapadia, N. I. Badler, and S. Jörg, "The effect of posture and dynamics on the perception of emotion," in *Proc. ACM Symp. Appl. Perception*, 2013, pp. 91–98.

[40] S. A. Etemad and A. Arya, "Extracting movement, posture, and temporal style features from human motion," *Biol. Inspired Cognitive Archit.*, vol. 7, pp. 15–25, 2014.

[41] H. G. Wallbott, "Bodily expression of emotion," *Eur. J. Soc. Psychol.*, vol. 28, no. 6, pp. 879–896, 1998.

[42] J. Posner, J. A. Russell, and B. S. Peterson, "The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology," *Develop. Psychopathol.*, vol. 17, no. 3, pp. 715–734, 2005.

[43] M. Coulson, "Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence," *J. Nonverbal Behavior*, vol. 28, no. 2, pp. 117–139, 2004.

[44] C. L. Roether, L. Omlor, A. Christensen, and M. A. Giese, "Critical features for the perception of emotion from gait," *J. Vision*, vol. 9, no. 6, pp 1–32, 2009.

[45] M. P. Murray, R. C. Kory, and S. B. Sepic, "Walking patterns of normal women," *Arch. Physical Med. Rehabil.*, vol. 51, no. 11, pp. 637–650, 1970.

[46] G. Mather and L. Murdoch, "Gender discrimination in biological motion displays based on dynamic cues," *Proc. Royal Soc. London Ser. B, Biol. Sci.*, vol. 258, no. 1353, pp. 273–279, 1994.

[47] S. A. Etemad, A. Arya, and A. Parush, "Additivity in perception of affect from limb motion," *Neurosc. Lett.*, vol. 558, pp. 132–136, 2014.

[48] O. Khatib, L. Sentis, K. Park, and J. Warren, "Whole-body dynamic behavior and control of human-like robots," *Int. J. Humanoid Robot.*, vol. 1, no. 1, pp. 29–43, 2004.

[49] P. Kingston and M. Egerstedt, "Motion preference learning," in *Proc. IEEE Am. Control Conf.*, 2011, pp. 3819–3824.

[50] A. LaViers and M. Egerstedt, "Style based robotic motion," in *Proc. IEEE Am. Control Conf.*, 2012, pp. 4327–4332.

[51] A. Samadani, R. Gorbet, and D. Kulic, "Affective movement recognition based on generative and discriminative stochastic dynamic models," *IEEE Trans. Human-Mach. Syst.*, vol. 44, no. 4, pp. 454–467, Aug. 2014.

S. Ali Etemad (M'09) received the Ph.D. degree in electrical and computer engineering from Carleton University, Ottawa, ON, Canada, in 2014.

His research was focused on human motion analysis for interactive multimedia and human–computer interaction (HCI) applications. He is currently an active expert in the industry as well as an Adjunct Professor with the School of Information Technology, Carleton University. His work and research is in the field of machine intelligence, pattern recognition, data science, and HCI.

Ali Arya (M'99–SM'06) received the Ph.D. degree in computer engineering from the University of British Columbia, Vancouver, BC, Canada, in 2003.

He joined the School of Information Technology, Carleton University, Ottawa, ON, Canada, in 2006, where he is currently an Associate Professor of interactive multimedia and design. His major research interests include computer graphics and animation, multimedia systems, human–computer interaction, virtual worlds and collaborative environments, artificial intelligence, and digital art.