Emotion Control of Unstructured Dance Movements

Andreas Aristidou
Shandong Univ. & Univ. of Cyprus
a.aristidou@ieee.org

Qiong Zeng
Shandong University
qiong.zn@gmail.com

Efstathios Stavrakis
University of Cyprus
stathis@cs.ucy.ac.cy

KangKang Yin
Simon Fraser University
kkyin@sfu.ca

Daniel Cohen-Or
Tel-Aviv University
cohenor@gmail.com

Yiorgos Chrysanthou
University of Cyprus
yiorgos@cs.ucy.ac.cy

Baoquan Chen
Shandong University & Advanced
Innovation Center for Future Visual
Entertainment
baoquan.chen@gmail.com

ABSTRACT

Motion capture technology has enabled the acquisition of high quality human motions for animating digital characters with extremely high fidelity. However, despite all the advances in motion editing and synthesis, it remains an open problem to modify pre-captured motions that are highly expressive, such as contemporary dances, for stylization and emotionalization. In this work, we present a novel approach for stylizing such motions by using emotion coordinates defined by the Russell's Circumplex Model (RCM). We extract and analyze a large set of body and motion features, based on the Laban Movement Analysis (LMA), and choose the effective and consistent features for characterizing emotions of motions. These features provide a mechanism not only for deriving the emotion coordinates of a newly input motion, but also for stylizing the motion to express a different emotion without having to reference the training data. Such decoupling of the training data and new input motions eliminates the necessity of manual processing and motion registration. We implement the two-way mapping between the motion features and emotion coordinates through Radial Basis Function (RBF) regression and interpolation, which can stylize freestyle highly dynamic dance movements at interactive rates. Our results and user studies demonstrate the effectiveness of the stylization framework with a variety of dance movements exhibiting a diverse set of emotions.

CCS CONCEPTS

• **Computing methodologies** → **Motion processing**; Motion capture; • **Information systems** → *Sentiment analysis*;

KEYWORDS

Computer Graphics, character animation, data-driven motion style transfer, motion editing, motion synthesis.

1 INTRODUCTION

Motion capture has been widely adopted for animating 3D digital characters due to the high quality of captured movements from human subjects. Once captured, however, the raw data contains neither semantics nor parameterization that may support further editing and control of the animations. In particular, human movements are characterized by style, which portrays, among others, its unique identity reflecting nuances of the character's emotion. Altering the motion styles is important for realistic virtual characters in computer games and animated films. The users, typically animators and directors, need to effectively control the style and emotion of a given animation, so as to better convey the story. Applying such stylistic motion editing remains a challenging problem, however, currently requiring expensive, time-consuming, and tedious manual interventions.

Previous studies on style transfer have mainly focused on locomotion, using data-driven models. Typically styles are learned and transferred from one animation clip to another that is carrying out the same motion task [Brand and Hertzmann 2000; Hsu et al. 2005]. More recently, Yumer and Mitra [2016] introduced a method based on spectral analysis that can be applied to different actions. However, motion registration is still required to align example clips performing structurally similar motion tasks. Yet it is not obvious as to how to perform such manual processing, including motion segmentation and registration, for unstructured movements.

In this paper, we focus on complex and highly dynamic contemporary dance performances. Such dances are more challenging to analyze than mundane tasks such as locomotion, as the performers can express their emotions through a rich motion repertoire.

Figure 1: Our framework stylizes highly dynamic movements with emotion variations. The first row shows a captured contemporary dance performance. The second row shows a "happier" dance, while the bottom row feels "sadder". In the happy motion the arms are more elevated; while in the sad dance the arms and torso are more depressed and bent.

Modern dances rebel against traditional ones by rejecting rigid rules and structures to favor freedom and expression of inner feelings. Therefore such performances contain a large variety of free style elements such as jumping, turning, and arm movements, and can express a richer emotion vocabulary beyond the capability of alternative performing arts. Therefore, motion segmentation and registration is hard to achieve, due to this lack of rules and structures. Moreover, no two modern dancers perform the same as there is no routine steps, elements or styles. Such motion irregularity and variety pose serious challenges to existing style editing techniques.

Our framework starts with extracting relevant spatiotemporal features from dance movements of known emotions, following the Laban Movement Analysis (LMA) system [Laban and Ullmann 2011]. Then a small set of effective and consistent features for emotion characterization are identified. These features are then used to map a new input motion to their emotion coordinates on the Russell's Circumplex Model (RCM) of affect [Russell 1980]. RCM is a widely-accepted model that distributes emotions in a two-dimensional circular space, with arousal and valence as its dimensions. The use of effective features and emotion coordinates decouples the training data and the input motion. That is, the training and testing data do not need to relate to each other structurally or semantically. Moreover, the training data themselves can be heterogeneous. The difficulties on motion segmentation and registration for contemporary dance movements are thus circumvented. Lastly, we map the required changes in emotion back to changes in motion features, and solve for a motion that best satisfies the desired features.

Our results demonstrate that our method is effective on emotion control of irregular motions, as shown in Figure 1. The main contributions of this work include:

(1) We are the first to analyze and control emotions for unstructured motions with an intuitive interface based on the RCM diagram, without the need for manual processing and registration of the training and input data.
(2) We present a simple statistic method to identify performer-independent consistent LMA motion features that are effective for emotion expression.
(3) We utilize RBF (Radial Basis Function) regression and interpolation for two-way mapping between motion features and RCM emotion coordinates. This enables stylizing motions with emotions in the full RCM space more than the discrete set of captured emotions.

2 RELATED WORK

Despite recent advances in motion analysis and parameterization [Kovar and Gleicher 2004; Krüger et al. 2010; Müller et al. 2005], developing methods for style editing and control remains challenging. Early representative methods used signal processing techniques [Bruderlin and Williams 1995; Unuma et al. 1995] and motion warping [Witkin and Popovic 1995]. Later, linear methods, such as multilinear models, were investigated to decompose motion data into action parameters and performer-dependent motion signatures [Vasilescu 2002], or extract "style" and "identity" variations to support locomotion editing [Min et al. 2010]. Similarly, Independent Component Analysis (ICA) was employed to separate motion data into visually meaningful components related to motion styles [Shapiro et al. 2006]. Physics-based methods such as [da Silva et al. 2008; Liu et al. 2005; Yin et al. 2008] can also generate multiple motion styles through tuning objective functions in optimization and control parameters in simulation. These methods are usually

hard to generalize to large-scale heterogeneous data and provide no intuitive control over the synthesized styles.

With more and more motion capture data available, nonlinear data-driven methods become the modern choice for generating stylistic variations of motions. Brand and Hertzmann [2000] trained generative Hidden Markov Models from unsupervised training data for synthesizing stylistic variations. A statistical motion model, called Linear Time Invariant (LTI) model, was proposed by Hsu *et al.* [2005] to encode style differences and transfer new input motion from one style to another. Gaussian Process models have also been used to capture stylistic variations of human movement for editing motions towards desired styles [Ikemoto et al. 2009; Wang et al. 2007]. More recently, Xia *et al.* [2015] proposed local mixtures of autoregressive models and an online algorithm to learn correspondences between styles and motions and perform style transfer. Most of these methods work only for motions performing similar tasks, require some degree of data processing and motion registration, and rely on subjective words from users for directing style editing. In contrast, our method can be applied to independent actions different from the training data. Our emotion editing interface based on the RCM diagram is more intuitive and principled to use.

Our motion analysis is based on the Laban Movement Analysis system [Laban and Ullmann 2011], which incorporates contributions from anatomy, kinesiology and psychology to describe, interpret and document human movements. The LMA principles have been widely used for motion analysis and synthesis [Chi et al. 2000; Hartmann et al. 2005; Luo and Neff 2012; Torresani et al. 2006; Zhao and Badler 2005]. More recently LMA features have also been used for motion retrieval and emotion classification [Aristidou et al. 2015; Kapadia et al. 2013; Senecal et al. 2016]. In particular, Aristidou *et al.* [2015] used a variety of LMA features to classify dance movements into discrete emotion classes using Random Forests. We extend this idea to regress motions to the continuous 2D space of RCM, and then further support motion synthesis to generate motions of user-desired emotions. One key component is to identify from all LMA features a subset of features that are effective and consistent in expressing emotions. These selected features serve as the bridge between motions and emotions.

Most recently, Yumer and Mitra [2016] introduced a method to transfer style between independent actions using spectral analysis. A pair of *source* and *target* motions performing the same action are chosen to extract their difference in the spectral domain, which is then applied to a new input motion that can contain an unseen action. In contrast, our method does not require selecting or registering specific motions from the training set, and the desired emotions can be specified directly on the RCM diagram. Automatic registration techniques have been proposed before [Chen et al. 2009], but mainly tested on locomotion. It would be quite challenging if possible at all to register freeform dances, as it is not obvious even to human users how the motion segments should align.

3 MOTION ANALYSIS

Analyzing motion characteristics that facilitate parameterized rather than data-driven style control is challenging. Our motion analysis algorithm encodes the stylistic characteristics of a training set
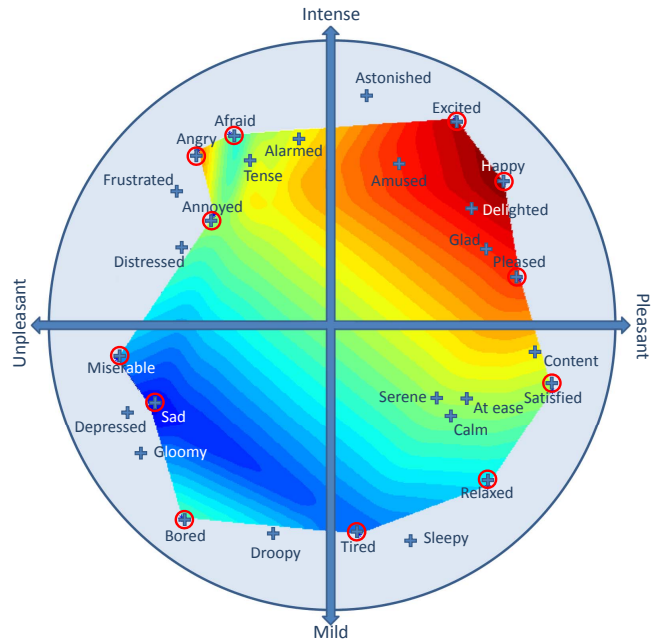

Figure 2: The Russell Circumplex Model (RCM) has two axes. The arousal axis spans from mild to intense emotions, and the valence axis spans from negative to positive emotions. We circled the twelve chosen emotions that the dancers were asked to perform on the RCM diagram. The colormap, or RCMF (RCM Function), shows the RBF interpolation for a specific feature. In this particular example, the chosen feature has the lowest value at the "Sad" motion, and the highest value at the "Happy" motion.

of dancing motions into higher-level features, which help decouple emotion control from the specific training motions that these features were learned from.

3.1 Data Collection and Feature Extraction

We learn high-level motion features based on the Laban Movement Analysis (LMA) principles, drawing from the framework described in [Aristidou et al. 2015]. Our dataset consists of 108 different contemporary dance performances captured at $480Hz$ and then downsampled to $30Hz$, each lasting approximately twenty seconds. Nine dancers were instructed to perform twelve dances in different emotional styles. The chosen emotions widely span the arousal and valence axes of the Russell's Circumplex Model of affect, as illustrated in Figure 2. We chose three emotions from each quadrant of the RCM diagram for the dancers to perform. Six dancers were asked to express emotions by following the same overall motion trajectory, while the remaining three dancers could act out their emotions completely freely. All captured dance motions are therefore labeled according to the designated emotion the performers were asked to express. We further retarget all motions to a single 3D skeleton with standard human proportions and body structure [Aristidou et al. 2015].
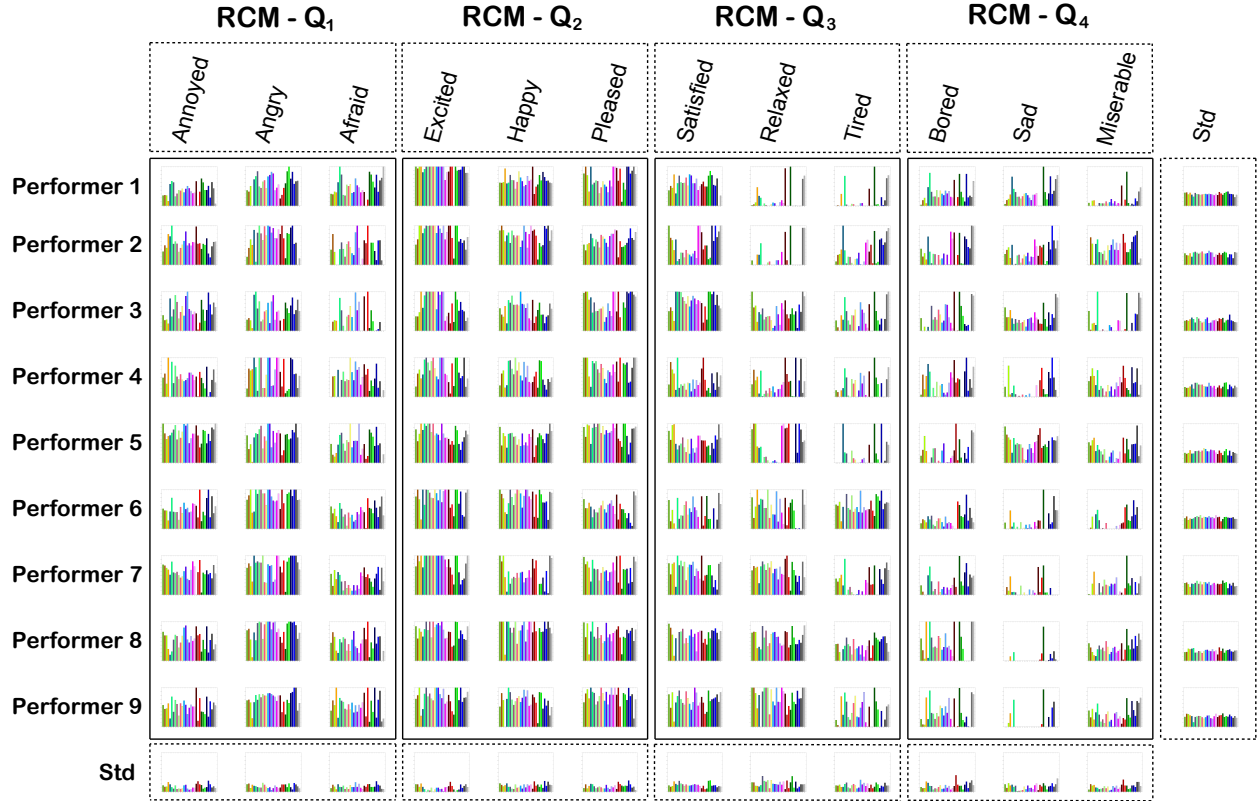
Figure 3: Visualization of the mean of the selected *effective* and *consistent* features for all the dancers and emotions, for one selected postural cluster. Each cell shows the mean feature value of a single motion. The last row is the standard deviation of the features across all dancers. The last column is the standard deviation of the features across all emotions.

We extract 37 basic spatio-temporal features $f = \{f^i\}, i = 1, \cdots, 37$ of human movements, which correspond to four Laban Movement Analysis components (BODY, EFFORT, SHAPE, SPACE). We further compute the minimum, maximum, mean and standard deviation of these features and derive 121 features $\tilde{f} = \{\tilde{f}^i\}, i = 1, \cdots, 121$ as shown in Table 1. Note that these derived features $\tilde{f}^i$ are normalized into the range $[0, 1]$, and averaged over a 35-frame sliding window to reduce noise. Hereafter we tend to omit the superscript of the feature values where possible for notation simplification.

3.2 Effective and Consistent Features

We wish to identify the emotion of a motion by analyzing the derived features $\tilde{f}$, but not all of them correlate with emotions. The subset of features that have distinct values for motions in different emotions is considered *effective* features that can help discern emotions parameterized by the RCM. Among the effective features, we also need to identify *consistent* features shared by multiple performers to express the same feelings. Hereafter we denote such *effective* and *consistent* features as selected features $\hat{f}$.

More specifically, we denote the mean value of a selected feature $\tilde{f}$ of a motion m as $\bar{f}$, and indicate its dancer p and expressed emotion e in parenthesis as $\bar{f}(p, e)$. To evaluate the effectiveness of

$\tilde{f}$, we examine the standard deviation of $\bar{f}(*, p)$ across emotions for a specific p. If $\mathrm{std}(\bar{f}(*, p)) \geq 0.21$, i.e., the variation of this feature across emotions is sufficiently large, then the feature is considered *effective*, but only for this particular performer p. If a feature is effective for the majority of the performers, e.g., seven out of nine performers, it is then counted as an overall effective feature.

To evaluate the consistency of $\tilde{f}$, we examine the standard deviation of $\bar{f}(e, *)$ across performers for a specific emotion e. If $\mathrm{std}(\bar{f}(e, *)) \leq 0.2$, i.e., the variation of this feature across performers is sufficiently small, then the feature is considered *consistent*, but only for this particular emotion e. If a feature is consistent for the majority of the emotions, e.g., ten out of twelve emotions, it is then counted as an overall consistent feature.

From all the derived features $\tilde{f}$, we identified 31 effective and consistent features as described above and denote them as selected features $\hat{f} = \{\hat{f}^i\}, i = 1, \cdots, 31$ hereafter. These features are highlighted in orange in Table 1. We further visualize these features in Figure 3. The last column of Figure 3 shows the standard deviation of features across emotions, and high values indicate the effectiveness; while the last row shows the standard deviation of features across performers, and low values indicate the consistency. Figure 4 further shows the RCM diagram of each performer for one selected feature.

		Basic LMA Features f^i	Derived Features $\tilde{f}^i$			
		Description	max	min	std	$mean$
BODY	f^1	Left foot-hip distance	$\tilde{f}^1$	$\tilde{f}^2$	$\tilde{f}^3$	$\hat{f}^4$
	f^2	Right foot-hip distance	$\tilde{f}^5$	$\tilde{f}^6$	$\tilde{f}^7$	$\hat{f}^8$
	f^3	Left hand-shoulder distance	$\tilde{f}^9$	$\tilde{f}^{10}$	$\tilde{f}^{11}$	$\hat{f}^{12}$
	f^4	Right hand-shoulder distance	$\tilde{f}^{13}$	$\tilde{f}^{14}$	$\tilde{f}^{15}$	$\hat{f}^{16}$
	f^5	Hands distance	$\tilde{f}^{17}$	$\tilde{f}^{18}$	$\tilde{f}^{19}$	$\tilde{f}^{20}; \hat{f}^1$
	f^6	Left hand-head distance	$\tilde{f}^{21}$	$\tilde{f}^{22}$	$\tilde{f}^{23}$	$\hat{f}^{24}$
	f^7	Right hand-head distance	$\tilde{f}^{25}$	$\tilde{f}^{26}$	$\tilde{f}^{27}$	$\hat{f}^{28}$
	f^8	Left hand-hip distance	$\tilde{f}^{29}$	$\tilde{f}^{30}$	$\tilde{f}^{31}$	$\tilde{f}^{32}; \hat{f}^2$
	f^9	Right hand-hip distance	$\tilde{f}^{33}$	$\tilde{f}^{34}$	$\tilde{f}^{35}$	$\tilde{f}^{36}; \hat{f}^3$
	f^{10}	Hip-ground distance	$\tilde{f}^{37}$	$\tilde{f}^{38}$	$\tilde{f}^{39}$	$\tilde{f}^{40}; \hat{f}^4$
	f^{11}	Hip-ground minus feet-hip	$\tilde{f}^{41}$	$\tilde{f}^{42}$	$\tilde{f}^{43}$	$\hat{f}^{44}$
	f^{12}	Feet distance	$\tilde{f}^{45}$	$\tilde{f}^{46}$	$\tilde{f}^{47}$	$\tilde{f}^{48}; \hat{f}^5$
	f^{13}	Left hand and chest	$\tilde{f}^{113}$	$\tilde{f}^{114}$	$\tilde{f}^{115}$	$\tilde{f}^{116}; \hat{f}^{30}$
	f^{14}	Right hand and chest	$\tilde{f}^{117}$	$\tilde{f}^{118}$	$\tilde{f}^{119}$	$\tilde{f}^{120}; \hat{f}^{31}$
EFFORT	f^{15}	Deceleration peaks				$\tilde{f}^{49}; \hat{f}^6$
	f^{16}	Pelvis velocity	$\tilde{f}^{50}$		$\tilde{f}^{51}$	$\tilde{f}^{52}; \hat{f}^7$
	f^{17}	Left hand velocity	$\tilde{f}^{53}$		$\tilde{f}^{54}$	$\tilde{f}^{55}; \hat{f}^8$
	f^{18}	Right hand velocity	$\tilde{f}^{56}$		$\tilde{f}^{57}$	$\tilde{f}^{58}; \hat{f}^9$
	f^{19}	Left foot velocity	$\tilde{f}^{59}$		$\tilde{f}^{60}$	$\tilde{f}^{61}; \hat{f}^{10}$
	f^{20}	Right foot velocity	$\tilde{f}^{61}$		$\tilde{f}^{61}$	$\tilde{f}^{64}; \hat{f}^{11}$
	f^{21}	Pelvis acceleration	$\tilde{f}^{65}; \hat{f}^{12}$		$\tilde{f}^{66}$	
	f^{22}	Left hand acceleration	$\tilde{f}^{67}; \hat{f}^{13}$		$\tilde{f}^{68}$	
	f^{23}	Right hand acceleration	$\tilde{f}^{69}; \hat{f}^{14}$		$\tilde{f}^{70}$	
	f^{24}	Left foot acceleration	$\tilde{f}^{71}; \hat{f}^{15}$		$\tilde{f}^{72}$	
	f^{25}	Right foot acceleration	$\tilde{f}^{73}; \hat{f}^{16}$		$\tilde{f}^{74}$	
	f^{26}	Jerk	$\tilde{f}^{75}; \hat{f}^{17}$		$\tilde{f}^{76}$	
SHAPE	f^{27}	Volume (5 joints)	$\tilde{f}^{77}$	$\tilde{f}^{78}$	$\tilde{f}^{79}$	$\tilde{f}^{80}; \hat{f}^{18}$
	f^{28}	Volume (All joints)	$\tilde{f}^{81}$	$\tilde{f}^{82}$	$\tilde{f}^{83}$	$\tilde{f}^{84}; \hat{f}^{19}$
	f^{29}	Torso height	$\tilde{f}^{85}$	$\tilde{f}^{86}$	$\tilde{f}^{87}$	$\tilde{f}^{88}; \hat{f}^{20}$
	f^{30}	Hands level				$\tilde{f}^{89}-\tilde{f}^{91}; \hat{f}^{21}-\hat{f}^{23}$
	f^{31}	Volume (upper body)	$\tilde{f}^{97}$	$\tilde{f}^{98}$	$\tilde{f}^{99}$	$\tilde{f}^{100}; \hat{f}^{26}$
	f^{32}	Volume (lower body)	$\tilde{f}^{101}$	$\tilde{f}^{102}$	$\tilde{f}^{103}$	$\tilde{f}^{104}; \hat{f}^{27}$
	f^{33}	Volume (right side)	$\tilde{f}^{105}$	$\tilde{f}^{106}$	$\tilde{f}^{107}$	$\tilde{f}^{108}; \hat{f}^{28}$
	f^{34}	Volume (left side)	$\tilde{f}^{109}$	$\tilde{f}^{110}$	$\tilde{f}^{111}$	$\tilde{f}^{112}; \hat{f}^{29}$
SPACE	f^{35}	Total distance				$\tilde{f}^{92}; \hat{f}^{24}$
	f^{36}	Area per second	$\tilde{f}^{93}$	$\tilde{f}^{94}$	$\tilde{f}^{95}$	$\tilde{f}^{96}; \hat{f}^{25}$
	f^{37}	Total volume				$\hat{f}^{121}$

Table 1: The 37 basic LMA features f^i grouped by LMA component (BODY, EFFORT, SHAPE, SPACE), the 121 derived features $\tilde{f}^i$, and the 31 selected features $\hat{f}^i$ which are *effective* and *consistent*. Note that each postural cluster only uses a subset of these selected features.

3.3 Mapping Motion to Emotion

In order to analyze the emotion of a new input motion, we need to parameterize the motion with the RCM emotion coordinates $e = (e_x, e_y)$, through the selected feature vector $\hat{f}$. We use Gaussian Radial Basis Function (RBF) regression analysis to model the correlation between motion features and emotion coordinates [Rasmussen and Williams 2005]. More specifically, to estimate e_x of a new motion with feature $\hat{f}$:

$$e_x = w_0 + \sum_{i=1}^{31} w_i \hat{f}^i + \sum_{k=1}^{12} \lambda_k \phi(||\hat{f} - \hat{f}_k||), \qquad (1)$$

and

$$\phi(r) = \exp(-\frac{r^2}{2\sigma^2}), \qquad (2)$$

where σ is the average distance between the selected features of the captured motions. The weights w_0, w_i, λ_k are obtained by fitting Equation 1 to the features of the captured emotions $\hat{f}_k$. Also note that we have multiple dancers performing the same emotion, $\hat{f}_k$ is actually the mean value across all performers for a specific emotion
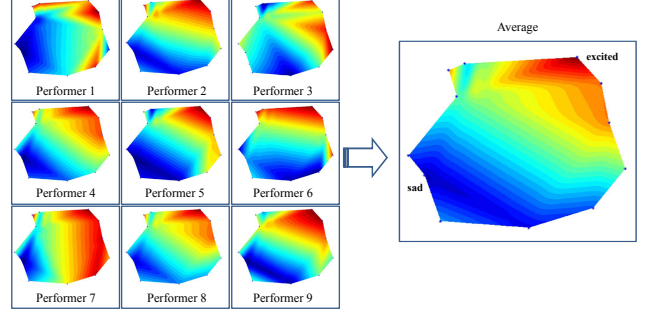

Figure 4: Visualization of the RCM Functions of a selected feature: the velocity of the left hand. High values are marked with red and low values with blue. It can be observed that the feature has a consistent gradient across all nine performers shown on the left. The average RCMF of all performers is shown on the right. In this example, the velocity of the left hand has a high mean value when the emotion is "excited" and a low mean value when the emotion is "sad".

e_k. We can perform a similar regression on e_y as well. We choose to use regression rather than classification, as we wish to support emotions varing continuously over the RCM space. Our problem is different from traditional emotion recognition problems. In motion editing for emotion control, a motion can have a mixture of characteristics of multiple emotions, and motions change continuously so do the emotions expressed by the motions.

We use leave-one-out cross validation to test the accuracy of the regression. We use 96 out of the 108 captured motions, or the performance of eight out of the nine dancers, as the training set, and the remaining twelve motions as the testing set. We rotate the dancers to repeat such regression nine times, each using a different performer as the testing set. We further employ Principal Component Analysis (PCA) to reduce the dimensionality of the selected features from 31 to 20, which covers 95% of the feature variance. The accuracy of the regression, as measured by the distance to the desired RCM coordinates, is improved by 15% after PCA. The results are shown in Figure 5. As we can see, the emotion coordinates do not really align with each other or the original position from the RCM model marked by the blue dots. This implies the complexity of emotions. However, all the predicted emotion coordinates do fall into the correct quadrant of the RCM diagram. Note that the emotion coordinates predicted for a new motion are clamped to be within the RCM diagram, as necessary for one of the bored motions in Figure 5.

3.4 Mapping Emotion to Motion

To edit and control the emotion expressed in a given motion, we need to map changes in emotion coordinates back to the space of motion features. Again, we use RBFs to estimate the feature value $\hat{f}^i$ at an arbitrary emotional coordinate e on the RCM diagram:

$$\hat{f}^i = w_0 + w_1 e_x + w_2 e_y + \sum_{k=1}^{12} \lambda_k \phi(||e - e_k||). \qquad (3)$$
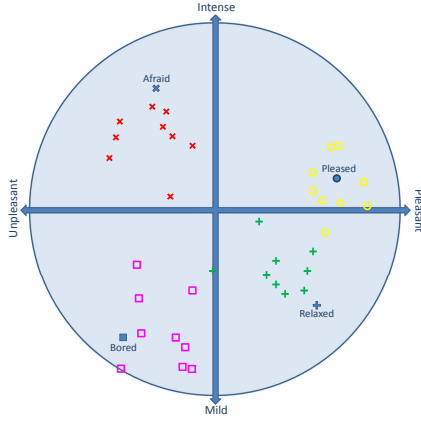
Figure 5: Leave-one-out cross validation of the Gaussian RBF regression. Predictions of "Afraid" are in red, "Bored" in magenta, "Relaxed" in green, and "Pleased" in yellow. Each of the nine predictions per emotion is obtained by a different regression analysis, with motions from one dancer as the testing data and the motions from the remaining eight dancers as the training set. The positions of the blue dots are determined by the original RCM model [Russell 1980]

We use multiquadrics radial basis functions:

$$\phi(r) = \sqrt{1 + \frac{r^2}{\sigma^2}}, \qquad (4)$$

where σ is the average distance between the emotion coordinates of the captured motions. Again, the weights w_0, w_1, w_2, λ_k are fitted to the captured motions.

Figure 2 visualizes one feature value interpolated from the captured RBF centers e_k over the RCM diagram, and we will refer to these continuous spaces RCM Functions (RCMFs). RCMFs enable users to go beyond the twelve chosen discrete emotions and specify a desired emotion anywhere inbetween on the RCM diagram.

3.5 Cluster Based Feature Analysis

Contemporary dances are free-form motions that contain extreme postures. These postures can produce salient features that may or may not be valid for other postures. For example, some performers crouch the whole body down to the floor to express the sad emotion. Lowering the height of the root can thus be picked up as an effective feature to express sadness by the feature analysis step described in Section 3.2. However, applying this feature blindly to other poses, such as walking stances, can result in walking with bent knees, which is strenuous and tiring and not desirable by a sad performer. Therefore it is necessary to identify and apply salient features only for poses that are relatively nearby. We thus cluster all the training poses and then perform local feature analysis within each cluster to alleviate such faulty generalization of motion features.

We cluster all the captured poses using hierarchical clustering [Manning et al. 2008]. Our character model has one root joint and 20 internal joints. To measure the similarity between poses, we first discard the planar position and facing direction of the root joint, so that each pose is parameterized by the height of the root

Feature Modification Rules	Affected Features
$g_1(c_1)$: modifies the position of the pelvis p_r: if $c_1 > 1.0$ if $\|d_{pelvis}\| > \|d_{pelvis,T-pose}\|$ $\quad p_r += (d_{pelvis} - d_{pelvis,T}) \times c_1$ else $\quad p_r += (d_{pelvis,T} - d_{pelvis})(1.0 - 1.0/c_1)$ else $\quad p_r += d_{pelvis} \times (c_1 - 1.0)$ where d_{pelvis} is the vector from the pelvis to the ground, and $d_{pelvis,T}$ is the same vector in the T-pose.	$\hat{f}^4, \hat{f}^{18}, \hat{f}^{19}, \hat{f}^{27}, \hat{f}^{28},$ $\hat{f}^{29}$.
$g_2(c_2)$: modifies the position of the head p_h: $p_h += d_{torso,T} - [(d_{torso,T} - d_{torso})/c_2] - d_{torso}$ where d_{torso} is the vector from the head to the pelvis, and $d_{torso,T}$ is the same vector in the T-pose.	$\hat{f}^{18}, \hat{f}^{19}, \hat{f}^{20}, \hat{f}^{26}, \hat{f}^{28},$ $\hat{f}^{29}$.
$g_3(c_{L/R}^{head}, c_{L/R}^{chest}, c_{L/R}^{pelvis}, c_{L/R}^{ground})$: modifies the position $p_{L/R}$ of the left (L) and right (R) hand respectively: $p_{L/R} += [d_{L/R}^{head} \times (c_{L/R}^{head} - 1.0) + d_{L/R}^{chest} \times (c_{L/R}^{chest} - 1.0) +$ $d_{L/R}^{pelvis} \times (c_{L/R}^{pelvis} - 1.0) + d_{L/R}^{ground} \times (c_{L/R}^{ground} - 1.0)]/4$ where $d_{L/R}^{joint}$ is the vector from each hand to the referred joint (head, chest, pelvis, ground).	$\hat{f}^1, \hat{f}^2, \hat{f}^3, \hat{f}^{18}, \hat{f}^{19}, \hat{f}^{21},$ $\hat{f}^{22}, \hat{f}^{23}, \hat{f}^{26}, \hat{f}^{28}, \hat{f}^{29},$ $\hat{f}^{30}, \hat{f}^{31}$.
$g_4(c_4)$: modifies the frame rate ν: $\nu \times= c_4$	$\hat{f}^6, \hat{f}^7, \hat{f}^8, \hat{f}^9, \hat{f}^{10}, \hat{f}^{11},$ $\hat{f}^{12}, \hat{f}^{13}, \hat{f}^{14}, \hat{f}^{15}, \hat{f}^{16},$ $\hat{f}^{17}, \hat{f}^{24}, \hat{f}^{25}$.

Table 2: Heuristic rules used to modify features for motion synthesis. Each rule is parameterized by one or more coefficients c_* and affects multiple features.

and 21 quaternions $p = \{h, q_j\}$. We then compute the difference between pose p_u and p_v as follows [Lee et al. 2002]:

$$d(p_u, p_v) = (h_u - h_v) + \sum_j w_j || \log(q_{u_j}^{-1} q_{v_j})||^2 \qquad (5)$$

We observe that the lower-body joints are mainly responsible for balance and locomotion, and emotions are mainly expressed through upper-body movements. Hence we assign higher weights to the root and arm joints in comparing poses. We experimented with different numbers of clusters, and found that using five clusters represents a good trade-off between algorithm effectiveness and algorithm complexity. These clusters roughly classify all the poses into one crouching cluster of low stance, mid-air poses due to jumping, and three standing clusters with different arm postures.

We identify effective and consistent features as described in Section 3.2 locally in each of the five clusters. For each cluster, we calculate the varying spatiotemporal features of the performances for each dancer and each emotion to generate a cluster-dependent emotion-content table. All the five local emotion-content tables can be found in the supplementary materials.

4 MOTION SYNTHESIS

Given an input motion m, we first analyze its emotion e from its motion features as described in Section 3.3. We then visualize e on the RCM diagram and then the user is free to indicate a new desired emotion e^* to transform the input motion to. Our motion synthesis algorithm to be described then outputs a new motion m^*. The user can perform the editing multiple times until m^* is satisfactory. Note that we apply motion synthesis only on sampled keyframes, i.e., every fifth frame uniformly drawn from the motion. We then interpolate the edited keyframes to obtain the full motion

Figure 6: Modification of an input motion towards a sadder (top blue) and happier (bottom red) emotion. Motions are synchronized in time for ease of comparison. The sad character draws its limbs and torso closer to the body; while the happy character extends and stretches more.

trajectory. Therefore what follows is the description of the keyframe synthesis.

We first map the desired emotion e^* to desired motion features $\hat{f}^*$ as described in Section 3.4. As the features $\hat{f}^{*i}$ are not guaranteed to be independent of or even consistent with each other, e.g., the desired hand-hand distance may be conflicting with the desired hand-hip distance, we first optimize for poses that satisfy the desired features as best as possible. To this end, we designed four heuristic rules as shown in Table 2 to modify the positions of four control joints: head, hands, and pelvis. For example, rule $g_1(c_1)$ modifies the position of the pelvis joint, which affects the hip-ground distance $\hat{f}^4$, as well as volumes $\hat{f}^{17}, \hat{f}^{18}, \hat{f}^{27}, \hat{f}^{28}, \hat{f}^{29}$. We search for c_1, the coefficient for rule g_1, that minimizes the distance between the current feature values and the desired feature values for all the keyframes indexed by time t in a least-squares sense:

$$\min_{c_1} \sum_t ||\hat{f}_t^* - \hat{f}_t(c_{1_t})||^2 \qquad (6)$$

We use the interior-reflective Newton method [Coleman and Li. 1996] for this keyframe optimization and we initialize all coefficients to 1.0.

We then sequentially apply the rest of the rules g_2, g_3, g_4 to optimize for the coefficients of these rules to minimize the distance between the current feature and the desired feature. From these optimized coefficients we then compute the desired positions for all the control joints. We then pass these joint positions to an Inverse Kinematics solver, (FBBIK) [Root-Motion 2016], to solve for the fullbody poses. FBBIK automatically respects skeletal constraints and guarantees the naturalness of the reconstructed fullbody poses. This is important as the desired control joint positions from the optimizations above may violate physical constraints, such as requiring variable bone length or out of range motions. FBBIK also supports synergies between nearby joints on the same kinematic chain, by providing manually tunable *pull* weights.

5 RESULTS

We have implemented our emotion control framework using MAT-LAB R2014b. We report the performance on an eight-core PC with
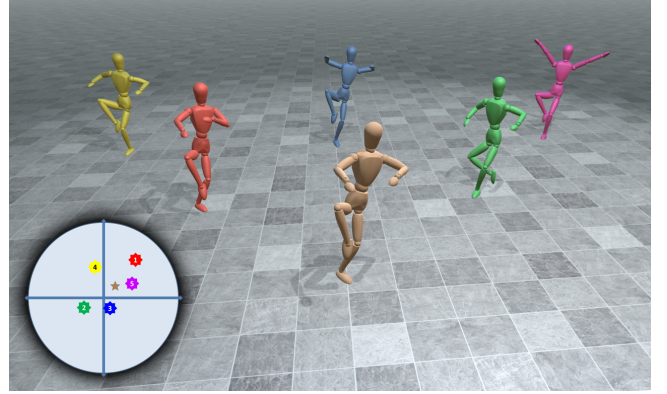

Figure 7: Six characters dancing with different emotions. The emotion coordinates are visualized in the lower-left inset with corresponding colors. The brown character shows the original motion. The two characters in the middle dance more happily (red) or more sadly (green). The three dances at the back are edited towards afraid (yellow), tired (blue), and pleased (purple), respectively.

Intel i7-2700K 3.5GHz CPU and 8GB RAM. Motion Analysis is performed offline, and takes around 45 minutes with our 108 dances. The motion synthesis is interactive but not realtime, and takes 12 seconds for an input motion of 680 frames. We have generated a variety of stylized dances by shifting the expressed emotions of new input motions using our framework. We encourage the readers to check the accompanying video for side-by-side comparisons of the input motions and the edited motions. Additional example results are also provided in the supplementary material.

Emotions are controlled by simply specifying the target emotion coordinates on the 2D RCM diagram for various input motions. We focus on generating natural motion variations [Lau et al. 2009] rather than completely reversing extreme emotions. In Figure 1 one motion is modified to express a happier and a sadder feeling. Figure 6 overlays the original motions in brown with the stylized motions in a different color, which are temporally synchronized for ease of comparison. The happier character elevates its limbs and torso more in comparison to the input motion; while the sadder character lowers its limbs and bends the torso more. In addition, the happy character moves faster and the sad character slows down, which can be best observed in the video. In Figure 7, we show an example frame of a group dance where each dancer performs in different emotion. The brown character performs the original dance, while each of the remaining five characters are stylized with a different target emotion, including happy, sad, afraid, tired, and pleased.

We have also tested our method on stylizing non-dance motions, such as locomotion including walking and running obtained from the CMU motion capture database. Our stylizer can be applied without any additional data-dependent processes. However, the mapping of such motion onto the RCM diagram is less accurate due to the different nature of the training and testing data. Nevertheless, experiments show that our framework can still generate reasonable emotional variants of the input motion. Figure 8 illustrates an
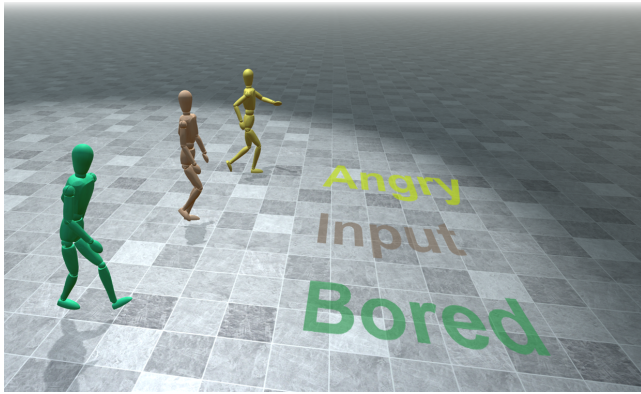
Figure 8: Stylization on walking. The input motion is visualized on the brown character, whereas the modified motion towards angry and bored is shown on the yellow and green characters, respectively.

Ground Truth \ User Response	More Pleasant	Nothing Changed	Less Pleasant
More Pleasant	75.8%	19.3%	4.9%
Less Pleasant	5.8%	15.3%	78.9%

Ground Truth \ User Response	More Intense	Nothing Changed	Less Intense
More Intense	78.4%	18.3%	3.2%
Less Intense	5.0%	14.4%	80.5%

Table 3: Emotion control along the RCM axes. More than 75% of the responses are correct along the valence axis, and more than 78% correct along the arousal axis.

example of modifying a walking sequence, where the brown character shows the input motion, and the yellow and green character perform a stylized angry and bored walking, respectively. Including generic motions into the training dataset may further improve the synthesized motion quality.

We validate our motion editing results by calculating the emotion coordinates of the modified dances back onto the RCM diagram again. We found that the new emotion coordinates are close to the desired ones even though they do not align exactly. This indicates that our method is effective in achieving desired emotionalization. Furthermore, we conducted three user studies to be described shortly to perceptually evaluate the effectiveness of our results. In each user study, the target emotion coordinates for the edited motions were on average 23% of the RCM diameter away from their original coordinates. Also participants with incomplete responses were discarded and were not counted in the evaluation. For more detailed information such as the questionnaires used in our user studies, please refer to our supplementary materials.

5.1 Naturalness

In the first user study, we investigate the naturalness of the edited motion with respect to the original motion capture data. 65 participants (28 males and 37 females, including 7 dance experts) were presented with six pairs of dance motions rendered on the 3D wooden mannequin character. The motion pairs all consist of one captured motion, and one edited motion by our algorithm. The two motions were placed side by side with a random left-right position. The participants were asked to point out the original unaltered motion for each pair, or "Do not know" if they cannot decide. The users picked the true raw motion for 48% of the motion pairs; they picked the stylized motion for 41% of the pairs; and for 11% of the pairs they could not decide. A paired t-test on the probabilities of picking the raw and the edited motion for the six pairs of motions returned $t = 1.8712$ and $p = 0.1202$, and no statistically significant difference was found. We thus conclude that the edited motion cannot be easily distinguished from raw motion capture data.

5.2 Effectiveness

In the second user study, we investigate the effectiveness of our emotion control framework along the RCM axes. 108 participants (40 males and 68 females, including 15 dance experts) were provided with a short description of the RCM diagram. Then they were shown eight motion pairs rendered on the same 3D wooden mannequin. The motion pairs all consist of one captured motion, and one edited motion. Participants were asked to answer how they think the edited motion differs from the original. Two sets of answers were provided, the first was about the valence of the motion and the options given were: "more pleasant", "nothing changed" and "less pleasant", and the second set of options was about the intensity of the motion and the options given were: "more intense", "nothing changed" and "less intense". Table 3 summarizes the responses of the participants. More than 75% of the emotion editing were correctly identified along the RCM valence axis, and more than 78% along the RCM arousal axis. This indicates that our emotion control along the RCM axes are effective as well as the intended direction of changes.

5.3 Recognition

In the third user study, we investigate the emotion recognition performance in the full RCM space. The same set of participants as in the Effectiveness User Study described above were used. A set of 12 motions were chosen to broadly distribute across all four RCM quadrants: RCM-Q_1,RCM-Q_2,RCM-Q_3,RCM-Q_4. Six motions were the captured motions and the remaining six were our edited motions. All participant were shown these same 12 motions and asked to assign an emotion for each of them. They were given five choices: one emotion from each RCM quadrant, and a fifth choice "none" in case they cannot decide. The ability to recognize the respective quadrant (RCM-Q_i) of the expressed emotion is reported in Table 4. The results show that the success rates of emotion recognition from raw motions and our edited motions are roughly the same. More specifically, the participants were able to identify the correct emotion from real mocap motions with an average accuracy of 68%, and 67% from our stylized motions.

6 CONCLUSIONS AND FUTURE WORK

We have presented an emotion control framework for unstructured dances. We have shown how to map freeform motion sequences

User Response ⟍ Ground Truth	RCM-Q_1	RCM-Q_2	RCM-Q_3	RCM-Q_4	None
RCM-Q_1	60.6%	18.3%	10.1%	6.4%	4.6%
RCM-Q_2	4.0%	74.0%	11.0%	6.7%	4.3%
RCM-Q_3	1.8%	0.9%	68.8%	28.4%	0.0%
RCM-Q_4	8.3%	0.0%	19.3%	68.8%	3.7%

(a) Emotion Recognition of Capture Motions

User Response ⟍ Ground Truth	RCM-Q_1	RCM-Q_2	RCM-Q_3	RCM-Q_4	None
RCM-Q_1	57.8%	11.0%	5.5%	6.4%	19.3%
RCM-Q_2	6.9%	75.2%	3.7%	6.0%	8.3%
RCM-Q_3	3.7%	1.8%	67.0%	16.5%	11.0%
RCM-Q_4	12.8%	1.8%	14.7%	67.9%	2.8%

(b) Emotion Recognition of Edited Motions

Table 4: Emotion recognition in the full RCM space from captured motions (a) and edited motions (b). RCM-Q_1 corresponds to the intense and unpleasant emotions; RCM-Q_2 the intense and pleasant; RCM-Q_3 the mild and pleasant; and RCM-Q_4 the mild and unpleasant. Each row represents the RCM quadrant that the emotion of the presented motion lies in; and each column shows the percentage of user responses. Participants had a similar performance in emotion recognition from captured motions and stylized motions.

to the 2D space of the Russell's Circumplex Model of affect; and how to synthesize altered motions with different desired emotions, usually interactively adjusted by users. This stylization framework does not require any manual motion preprocessing such as segmentation, labeling, or registration. Thus it can handle heterogenous and unstructured motions as well as regular movements. Three user studies verified the quality of our results and the effectiveness of the proposed framework.

We foresee this framework useful for content creators to generate multiple variations of captured or authored motions [McDonnell et al. 2008b]. The RCM interface is intuitive for both experts and novice users. The performance is interactive and likely realtime after code optimization. Training with application specific data or generic motions will likely affect the synthesis quality to some extent, and is worth exploring as a future direction.

The results of our motion analysis based on selected LMA features correspond well to findings from previous works on how emotions are typically expressed through body motions in dance [Normoyle et al. 2013; Sawada et al. 2003; Shiratori et al. 2006]. For example, higher arm motions convey a happier emotion, sad movements tend to be slower with the head and spine more bent, a lot of the emotion is conveyed through upper body motions. In the future, we would like to explore mid-level pose features [Pons-Moll et al. 2014] to see if they are more effective. Our feature selection based on variance is also simplistic, and could be improved by selecting features that discriminate between emotions and improve emotion prediction [McDonnell et al. 2008a; Sigal et al. 2010].

We have found that emotion recognition from free-form dances is in general a harder problem than we thought. For example, the accuracy of recognizing the RCM quadrant is only 68% for the captured motions, and is even lower if specific choices such as "excited" or "delighted" are demanded. However, we deem it the subjective nature of this problem: on one hand different motions are performed to express the same required emotions; on the other hand, emotions are recognized with different labels for the same motions. This is also testified by our regression results shown in Figure 5, where the predicted emotion coordinates for the captured motions spread out in each quadrant. Our observation also aligns well with findings from previous studies [Förger and Takala 2015]: users tend to agree more on the relative changes of emotions rather than their absolute labels. We did notice that professional dancers and choreographer usually perform slight better than non-expert users by 5-15% in the latter two user studies, due to their increased sensitivity to expressed emotions from their professional experiences.

We note that the quality of the stylized motions depends on the quality of the input data. In this work, we used an 8-camera motion capture system and 38 markers for each performer. Since contemporary dances often contain complex and extreme movements, such as turning around and crouching, marker occlusions occurred frequently. Therefore the quality of our captured motions were not great. In the future, we would like to use better motion capture system and capture setups. We also plan to integrate motion processing algorithms to remove data noise and better retargeting algorithms to remove foot sliding problems before passing the data to our motion stylizer.

Currently our motion synthesis method involves sequential optimizations on heuristic rules, and a third-party IK algorithm. In the near future we would like to reformulate it as one single constrained nonlinear optimization problem to achieve optimal results. We also wish to explore the latest advances in deep learning for feature selection and the two-way mapping between motions and emotions in the motion analysis part.

ACKNOWLEDGMENTS

REFERENCES

Andreas Aristidou, Panayiotis Charalambous, and Yiorgos Chrysanthou. 2015. Emotion analysis and classification: Understanding the performers emotions using the LMA entities. *Computer Graphics Forum* 34, 6 (Sept. 2015), 262–276.

Matthew Brand and Aaron Hertzmann. 2000. Style Machines. *ACM Trans. Graph.* (Sept. 2000), 183–192.

Armin Bruderlin and Lance Williams. 1995. Motion Signal Processing. In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '95)*. ACM, New York, NY, USA, 97–104.

Yen-Lin Chen, Jianyuan Min, and Jinxiang Chai. 2009. Flexible Registration of Human Motion Data with Parameterized Motion Models. In *Proceedings of the 2009 Symposium on Interactive 3D Graphics and Games (I3D '09)*. ACM, New York, NY, USA, 183–190.

Diane Chi, Monica Costa, Liwei Zhao, and Norman Badler. 2000. The EMOTE Model for Effort and Shape. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '00)*. New York, NY, USA, 173–182.

Thomas F. Coleman and Yuying Li. 1996. An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds. *SIAM J. Optimization* 6, 2 (1996),

418–445.

Marco da Silva, Yeuhi Abe, and Jovan Popović. 2008. Interactive Simulation of Stylized Human Locomotion. *ACM Trans. Graph.* 27, 3 (Aug. 2008), 82:1–82:10.

Klaus Förger and Tapio Takala. 2015. Animating with style: defining expressive semantics of motion. *The Visual Computer* (2015), 1–13.

Björn Hartmann, Maurizio Mancini, and Catherine Pelachaud. 2005. Implementing expressive gesture synthesis for embodied conversational agents. In *Proceedings of Gesture in Human-Computer Interaction and Simulation (GW'05)*. Springer-Verlag, 188–199.

Eugene Hsu, Kari Pulli, and Jovan Popović. 2005. Style Translation for Human Motion. *ACM Trans. Graph.* 24, 3 (July 2005), 1082–1089. https://doi.org/10.1145/1073204.1073315

Leslie Ikemoto, Okan Arikan, and David Forsyth. 2009. Generalizing Motion Edits with Gaussian Processes. *ACM Trans. Graph.* 28, 1 (Feb. 2009), 1:1–1:12.

Mubbasir Kapadia, I-kao Chiang, Tiju Thomas, Norman I. Badler, and Joseph T. Kider, Jr. 2013. Efficient motion retrieval in large motion databases. In *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games (I3D '13)*. ACM, NY, USA, 19–28.

Lucas Kovar and Michael Gleicher. 2004. Automated Extraction and Parameterization of Motions in Large Data Sets. *ACM Trans. Graph.* 23, 3 (Aug. 2004), 559–568.

Björn Krüger, Jochen Tautges, Andreas Weber, and Arno Zinke. 2010. Fast local and global similarity searches in large motion capture databases. In *Proc. of the ACM SIGGRAPH/Eurographics symposium on Computer animation (SCA'10)*. 1–10.

Rudolf Laban and Lisa Ullmann. 2011. The Mastery of Movement (4th ed.). Dance Book Ltd.

Manfred Lau, Ziv Bar-Joseph, and James Kuffner. 2009. Modeling spatial and temporal variation in motion data. *ACM Trans. on Graphics* 28, 5 (2009), Article 171.

Jehee Lee, Jinxiang Chai, Paul S. A. Reitsma, Jessica K. Hodgins, and Nancy S. Pollard. 2002. Interactive control of avatars animated with human motion data. *ACM Trans. on Graphics* 21, 3 (2002), 491–500.

C. Karen Liu, Aaron Hertzmann, and Zoran Popović. 2005. Learning Physics-based Motion Style with Nonlinear Inverse Optimization. *ACM Trans. Graph.* 24, 3 (July 2005), 1071–1081.

Pengcheng Luo and Michael Neff. 2012. A Perceptual Study of the Relationship between Posture and Gesture for Virtual Characters. In *Motion in Games*, Marcelo Kallmann and Kostas Bekris (Eds.). Lecture Notes in Computer Science, Vol. 7660. Springer Berlin Heidelberg, 254–265.

Christopher D. Manning, Prabhakar Raghavan, , and Hinrich Schütze. 2008. *Introduction to Information Retrieval*. Cambridge University Press.

Rachel McDonnell, Sophie Jörg, Joanna McHugh, Fiona Newell, and Carol O'Sullivan. 2008a. Evaluating the emotional content of human motions on real and virtual characters. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization*. ACM, 67–74.

Rachel McDonnell, Micheal Larkin, Simon Dobbyn, Steven Collins, and Carol O'Sullivan. 2008b. Clone attack! Perception of crowd variety. *ACM Trans. on Graphics* 27, 3 (2008), Article 26.

Jianyuan Min, Huajun Liu, and Jinxiang Chai. 2010. Synthesis and editing of personalized stylistic human motion. In *Proceedings of the 2010 ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games (I3D '10)*. ACM, NY, USA, 39–46.

Meinard Müller, Tido Röder, and Michael Clausen. 2005. Efficient content-based retrieval of motion capture data. *ACM Trans. on Graphics* 24, 3 (2005), 677–685.

Aline Normoyle, Fannie Liu, Mubbasir Kapadia, Norman I. Badler, and Sophie Jörg. 2013. The effect of posture and dynamics on the perception of emotion. In *Proc. of the ACM Symposium on Applied Perception*. 91–98.

Gerard Pons-Moll, David J. Fleet, and Bodo Rosenhahn. 2014. Posebits for Monocular Human Pose Estimation. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Carl Edward Rasmussen and Christopher K. I. Williams. 2005. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press.

Root-Motion. 2016. FINAL-IK: http://root-motion.com/, Accessed 10/2016. (2016).

James A. Russell. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology* 39 (1980), 1161–1178. Issue 6.

M Sawada, K Suda, and M Ishii. 2003. Expression of emotions in dance: relation between arm movement characteristics and emotion. *Preceptual and Motor Skills* 97, 3 (2003), 697–708.

Simon Senecal, Louis Cuel, Andreas Aristidou, and Nadia Magnenat-Thalmann. 2016. Continuous body emotion recognition system during theater performances. *Computer Animation and Virtual Worlds* 27, 3-4 (May 2016), 311–320.

Ari Shapiro, Yong Cao, and Petros Faloutsos. 2006. Style Components. In *Proceedings of Graphics Interface (GI '06)*. Canadian Information Processing Society, Toronto, Canada, 33–39.

Takaaki Shiratori, Atsushi Nakazawa, and Katsushi Ikeuchi. 2006. Dancing-to-Music Character Animation. *Computer Graphics Forum* 25, 3 (2006), 449–458.

Leonid Sigal, David J Fleet, Nikolaus F Troje, and Micha Livne. 2010. Human attributes from 3D pose tracking. In *European conference on computer vision*. Springer, 243–257.

Lorenzo Torresani, Peggy Hackney, and Christoph Bregler. 2006. Learning Motion Style Synthesis from Perceptual Observations. In *Proceedings of the Conference in Neural Information Processing Systems (NIPS '06)*. 1393–1400.

Munetoshi Unuma, Ken Anjyo, and Ryozo Takeuchi. 1995. Fourier Principles for Emotion-based Human Figure Animation. In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '95)*. ACM, New York, NY, USA, 91–96.

M. Alex O. Vasilescu. 2002. Human Motion Signatures: Analysis, Synthesis, Recognition. In *Proceedings of the 16th International Conference on Pattern Recognition (ICPR '02)*. IEEE Computer Society, Washington, DC, USA, 456–460.

Jack M. Wang, David J. Fleet, and Aaron Hertzmann. 2007. Multifactor Gaussian Process Models for Style-content Separation. In *Proceedings of the 24th International Conference on Machine Learning (ICML '07)*. ACM, New York, NY, USA, 975–982.

Andrew Witkin and Zoran Popovic. 1995. Motion Warping. In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '95)*. ACM, New York, NY, USA, 105–108.

Shihong Xia, Congyi Wang, Jinxiang Chai, and Jessica Hodgins. 2015. Realtime Style Transfer for Unlabeled Heterogeneous Human Motion. *ACM Trans. Graph.* 34, 4, Article 119 (July 2015), 10 pages. https://doi.org/10.1145/2766999

KangKang Yin, Stelian Coros, Philippe Beaudoin, and Michiel van de Panne. 2008. Continuation Methods for Adapting Simulated Skills. *ACM Trans. Graph.* 27, 3 (2008), Article 81.

M. Ersin Yumer and Niloy J. Mitra. 2016. Spectral Style Transfer for Human Motion Between Independent Actions. *ACM Trans. Graph.* 35, 4 (July 2016), 137:1–137:8.

Liwei Zhao and Norman I. Badler. 2005. Acquiring and validating motion qualities from live limb gestures. *Graphical Models* 67, 1 (2005), 1–16.