MODELING AND TRANSFORMATION OF 3D HUMAN MOTION

Seyed Ali Etemad
Department of Systems and Computer Engineering, Carleton University, 1125 Colonel By Drive, Ottawa, Canada

Ali Arya
School of Information Technology, Carleton University, 1125 Colonel By Drive, Ottawa, Canada

Keywords: Piece-wise Time Warping, Synthesis, Style Transformation, Critical Features.

Abstract: Applying different styles of motion such as those related to gender, age and energy of the performer, are important themes in creating realistic animation. While typically different sets of motion sequences must be created for different styles, in this paper we present a technique for transferring motion styles across different sequences. This allows transforming the style of a motion while preserving the primary action class of the original motion. A mathematical model capable of defining both the action class and the style class is proposed and based on the model, the conversion data are computed. A series of piece-wise time warping procedures are conducted prior to employing the defined transformation function. Using this technique, there would be no need for a new set of motion data to be captured or keyframed as the original motion data can be altered to show the desired style.

1 INTRODUCTION

The use of optical motion capture systems to produce high quality humanoid motion sequences has recently become increasingly popular in many areas such as high-end gaming, movie and animation production, and for educational purposes. Motion capture studios are expensive to employ or are unavailable to many people and locales. For this reason, the reuse of existing motion capture data and the resulting high-quality animation is an important area of research. Also strict motion capture data restricts users to pre-recorded movement that does not allow the addition of dynamic behaviours required in advanced games and interactive environment.

Procedural animation is creation of motion sequences through an algorithm rather than conventional methods like keyframing and motion capture. Creating new styles (i.e. personal variation of a base action) of animation based on existing motion sequences is an example and a relatively new topic of research in the field of animation where the goal is to manipulate existing sequences to create animation of the same action class, yet in a different mode. For instance using such techniques, a walking action sequence can be transformed from its initial masculine mode to a feminine one. Similarly the age and energy level of the characters can be controlled and the respective motion data can be synthesized.

The goal of this research is to create a system capable of learning various action classes with specific styles related to its character, and transform the style based on the user desire. Examples of the basic classes of action are the three periodic actions of walking, jumping and running. While walking and jumping are considerably different, walking and running are relatively similar, thus synthesis of motion data for these actions must be especially accurate.

For each basic class of action, secondary variation style pairs are necessary for adding a level of personality to the motions. Articles relating to personality and behaviour such as (Sok, et al., 2007; Su, et al., 2005; Kang & Kim 2007) were reviewed in order to determine contributing factors of human locomotion that can show a character's personality. These personality variations add a level of personality and interest to the basic, unexaggerated animation types mentioned earlier. Masculine to Feminine, Old to Young, Tired to Energetic, Happy to Sad, and Determined to Dreamy are these

Ali Etemad S. and Arya A. (2010).
MODELING AND TRANSFORMATION OF 3D HUMAN MOTION.
In *Proceedings of the International Conference on Computer Graphics Theory and Applications*, pages 307-315
DOI: 10.5220/0002755003070315

animation variation styles. By incorporating the variations with the basic animation types, new and unique character animations can be produced. The presented procedural animation method is a proof of concept production, and as such, only a selection of the animation types and variations are tested using our proposed method. The Walk, Run, and Jump animation types with the Masculine to Feminine, Old to Young, and Tired to Energetic variations are included. Yet the system can be extended to include other styles in the future.

We have proposed a novel mathematical model for describing human actions with various styles; the same model has been used to describe each step. The proposed transformation method is based on applying a transfer function obtained during a training phase to the base motion sequence in order to create a desired motion. The first step for style transformation of actions is for the training data to be temporally equal in length i.e. motion data matrices must have same number of frames in order for us to perform mathematical operations on them. We have proposed a novel piece wise time warping technique to convert our motion data sets to data sets of the same temporal length. The model has then been used to generate the necessary transfer functions for style transformations between different styles of the same action.

To test the outputs of our transformation techniques we formed a questionnaire and arranged for participants to exercise various transformations applied on different actions. The user comments and ratings confirm the significance of our research. Our procedural animation method offers animators high quality animations produced from an optical motion capture session, without incurring the cost of running their own sessions. This method utilizes a database of common animations sequences, derived from several motion capture sessions, which animators can manipulate and apply to their own existing characters through the use of our procedural animation technique.

2 RELATED WORK

In recent years, much research has been conducted with the aim of synthesizing human motion sequences. Statistical models have been one of the practical tools for human motion synthesis (Tanco & Hilton, 2000; Li, et al., 2002). Tanco and Hilton (2000, pp. 137-142) have trained a statistical model which employs a database of motion capture data for synthesizing realistic motion sequences and using the start and end of existing keyframes, original motion data are produced. Li et al. (2002, pp. 465-472) define a motion texture as a set of textons and their distribution values provided in a distribution matrix. The motion texton is modeled by a linear dynamic system (LDS). A maximum likelihood algorithm is designed to learn from a set of motion capture based textons. Finally, the learnt motion textures have been used to interactively edit motion sequences.

Egges et al. (2004, pp. 121-130) have employed principal component analysis (PCA) to synthesize human motion with the two deviations of small posture variations and change of balance. This approach is useful in cases where an animated character is in a stop/freeze situation where in reality no motionless character exists. Liu and Papovic (2002, pp. 408-416) have applied linear and angular momentum constraints to avoid computing muscle forces of the body for simple and rapid synthesis of human motion. Creating complex dynamic motion samples such as swinging and leaping have been carried out by Fang and Pollard (2003, pp. 417-426) using an optimization techniques applied along with a set of constraints, minimizing the objective function. Pullen and Bregler (2002, pp. 501-508) have trained a system that is capable of synthesizing motion sequences based on the key frames selected by the user. Their method employs the characteristic of correlation between different joint values to create the missing frames. In the end quadratic fit has been used to smooth the estimated values, resulting in more realistic looking results. Brand and Hertzmann (2000, pp. 183-192) employ probabilistic models for interpolation and extrapolation of different styles for synthesis of new stylistic dance sequences using a cross-entropy optimization structure which enables their style machine to learn from various style examples. Safonova et al. (2004, pp. 514-521) define an optimization problem for reducing the dimensionality of the feature space of a motion capture database, resulting in specific features. These features are then used to synthesize various motion sequences such as walk, run, jump and even several flips. This research shows that the complete feature space is not required for synthesis of human motion. We have employed this property in section 5 where correlated joints have been ignored when transforming the actor style themes.

Hsu et al. (2005, pp. 1082-1089) conduct style translations such as sideways walk and crouching walk based on a series of alignment mappings followed by space warping techniques using an LTI model. While this technique shows to be functional for the mentioned style translations, more minute style variations such as those related to gender, energy and age have not been tested for. Rose et al. (1998, pp. 32-40) have employed time warping as the first step towards synthesis of human motion which is an approach similar to ours. Specific kinematic constraints have been exercised along with interpolation of styles. The constraints, however, are selected manually and based on the nature of the action, as opposed to our proposed method in chapter 6 where we tend to automate the constraints. Time warping is also studied by Hsu et al. (2007, pp. 45-52) where they have used a specific reference motion for defining the constraints of the time warping procedure. Although this technique proves to be useful in some cases, it is not discussed whether the reference action is capable of determining the most suitable constraints for all cases.

3 DATASETS AND MOTION THEORY

Motion capture data obtained by means of a six MX40 camera Vicon system based in Carleton University have been used for this research. We have asked various actors to perform the required basic class of actions in each variation style, and the necessary database has been created.

The motion capture data come in the form of (1), where *Di* are the Cartesian values for the hip marker in 3D space with respect to the calibration origin and $\boldsymbol{\Theta_i}$ are the rotation angles in degrees for each marker. There are *m* rows, denoting *m* frames.

$$A = \begin{bmatrix} D_1 & , & \Theta_1 \\ D_2 & , & \Theta_2 \\ \vdots & & \vdots \\ D_m & , & \Theta_m \end{bmatrix} \tag{1}$$

In more details, the motion data matrices are in the form (2) where $\theta_i^{x_j}$ represents the rotation values of the x^{th} axis of the j^{th} marker for the i^{th} frame and d_i^x represents the position of the x^{th} axis of the hip marker for the i^{th} frame.

$$A = \begin{bmatrix} \left(d_1^x, d_1^y, d_1^z\right) & \left(\theta_1^{x_1}, \theta_1^{y_1}, \theta_1^{z_1}, \ldots, \theta_1^{x_n}, \theta_1^{y_n}, \theta_1^{z_n}\right) \\ \left(d_2^x, d_2^y, d_2^z\right) & \left(\theta_2^{x_1}, \theta_2^{y_1}, \theta_2^{z_1}, \ldots, \theta_2^{x_n}, \theta_2^{y_n}, \theta_2^{z_n}\right) \\ \vdots & \vdots \\ \left(d_m^x, d_m^y, d_m^z\right) & \left(\theta_m^{x_1}, \theta_m^{y_1}, \theta_m^{z_1}, \ldots, \theta_m^{x_n}, \theta_m^{y_n}, \theta_m^{z_n}\right) \end{bmatrix} \tag{2}$$

The frame rate of ***A*** is consistent throughout the whole matrix and the time lapse between each two consecutive frames is fixed and equal to 0.05 seconds. It is very important that all motion samples be of the same frame rate.

For this research, two different sets of data have been acquired. The first data set is used for training the system and creating the transformation values, while the second is used for testing the system as will be discussed later on in section 6.

Human action and motion in general is a combination of an action class and a set of stylistic variations adjoined to it. We call the action classes which are the dominant signals throughout the data sets, primary themes, and the stylistic variations as secondary motor themes (Hutchinson, 1996). Based on this definition, any action can be modeled by (3) where *Y[k,r]* is the action sequence with the primary theme *k* and secondary theme *r* as it is observed, *P[k]* is the primary motor theme of the same action class, and *S[r]* is the secondary theme of class *r*. In (3) *w[r]* is the weight applied to the secondary themes, and *e* represents the noise available in the model.

$$Y[k, r(1:f)] = P[k] + \sum_{r=1}^{f} \left(w[r] \cdot S[r]\right) + e \tag{3}$$

The model (3) is defined such that a combination of different styles can be applied to the same primary class of action. For instance, for action class jump, the secondary theme young-feminine can be defined. A total of *f* different *S* functions are foreseen in the model. Yet while training the system, all samples are designed and assumed to hold only one secondary theme. Each secondary theme function is learnt separately, and ultimately, they can be merged to form a multi-style secondary theme.

The goal of this research is to create the *S[r]* values, while trying to minimize *e*. In most cases, the *S[r]* signals are not powerful enough to influence the perception of *k*, i.e. the action class. But rarely, this scenario might take place. For instance if *r* relates to class of energy and speed related themes, *S[r]* along with a notable *w[r]* applied to a primary theme of walking can cause confusion as to if *Y[k,r]* was originally walking with a strong *S[r]* and *w[r]*, or whether *k* determines the action class of running for *Y[k,r]*, and *S[r]* and *w[r]* have been insignificant. This dilemma is addressed in section 6, yet further research is ongoing to tackle this issue.

4 FEATURE SELECTION AND PIECEWISE TIME WARPING

The captured motion data matrix $\boldsymbol{A}$ can be in various lengths. While some actions have been performed faster or slower due to actor preferences, some actions are by nature faster or slower than others. For instance, a 2-step walk cycle takes more time compared to a 2-step run cycle. In order to manipulate such data matrices based on one another, the temporal length of each action should be warped to match the temporal length of the relative actions with which we intend to blend. As for the number of columns in the matrices, they do not require any manipulation since the exact same number and orientation of markers have been used for all capture sessions.

The system is trained based on the original input sequence and the target sequence. The first set is the base animation while the second is the target sequence containing the desired secondary theme. Together they are called the training data. The system is then tested for new sets of action sequences called the test data. The training data along with the test sequence are temporally warped such that they are all in length. A straightforward approach is simple scaling; that is stretching or shrinking the motion sequences. Yet the sequences must be aligned appropriately. To tackle this problem, the warping of the matrices is carried out with the goal of aligning some critical features of the three datasets.

To conduct the very crucial piece-wise time warping, the signals are divided into two sections. The corresponding first sections are warped together, and the second parts are warped together as well. This is carried out in order for different corresponding sections of the action sequences to be aligned correctly. Thus, selecting the proper feature for which the signals are sliced at is very critical.

Different methods were tested for finding the critical features, used to determine the boundaries for the piece-wise time warping. The first method is by manually selecting the features. For walk and run sequences, the feature instance was selected as the occurrence of a foot touching the ground. This means in a sequence where walking starts with the right foot initiating the walk and ending in the same situation, the instance when the left foot touches the ground is labelled as the feature. For the jump sequence, the feature is defined as the instance where the actor is at the peak of his/her jump (when the inclining motion stops and returning to the ground initiates). Figures 1 and 2 illustrate the piecewise warping conducted using this technique. In Figure 1 through Figure 5, the data for the x axis of the marker placed on the right leg are plotted.

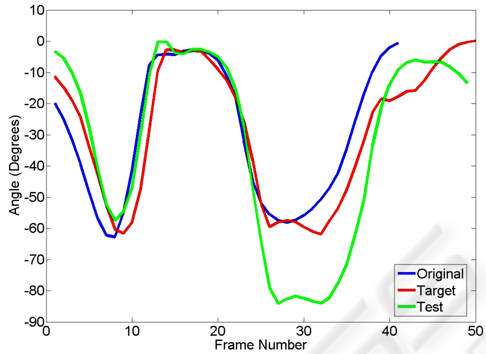

Figure 1: Signals from training, target, and test matrices before time warping.

It can be noticed from Figure 2 that the warped signals are stretched to last the exact amount of time (46 frames) which is the average of each of the individual signals in Figure 1. Yet misalignment in the global minimum peaks has occurred. Also the alignment of the signals between frame 20 and 30 is disposed of which is an undesired artefact. Thus selecting the features manually is not recommended.

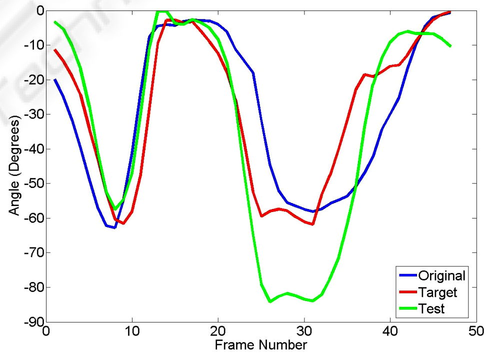

Figure 2: Signals from training, target, and test matrices after time warping using manual feature selection.

The second method is by means of statistical analysis. The $\boldsymbol{A}$ matrices are first normalized. The normalized version of $\boldsymbol{A}$ is called $\underline{\boldsymbol{A}}$. $\underline{\boldsymbol{A}}$ matrices are then differentiated with respect to time as presented by (4) where the result is denoted by $\vec{\underline{A}}$ presented by (5) and (6).

$$\vec{\underline{A}} = \frac{\partial \underline{A}}{\partial t} \tag{4}$$

$$\underline{\vec{A}}=\begin{bmatrix}\left(\theta_1^{x_1}-\theta_2^{x_1},\theta_1^{y_1}-\theta_2^{y_1},\theta_1^{z_1}-\theta_2^{x_n},\ldots,\theta_1^{x_n}-\theta_2^{x_n},\theta_1^{y_n}-\theta_2^{y_n},\theta_1^{z_n}-\theta_2^{z_n}\right)\\ \left(\theta_2^{x_1}-\theta_3^{x_1},\theta_2^{y_1}-\theta_3^{y_1},\theta_2^{z_1}-\theta_3^{z_1},\ldots,\theta_2^{x_n}-\theta_3^{x_n},\theta_2^{y_n}-\theta_3^{y_n},\theta_2^{z_n}-\theta_3^{z_n}\right)\\ \vdots\\ \left(\theta_{m-1}^{x_1}-\theta_m^{x_1},\theta_{m-1}^{y_1}-\theta_m^{y_1},\theta_{m-1}^{z_1}-\theta_m^{z_1},\ldots,\theta_{m-1}^{x_n}-\theta_m^{x_n},\theta_{m-1}^{y_n}-\theta_m^{y_n},\theta_{m-1}^{z_n}-\theta_m^{z_n}\right)\end{bmatrix} \tag{5}$$

$$\vec{A}=\begin{bmatrix}v_1^1,v_1^2,\cdots,v_1^{3n}\\ v_2^1,v_2^2,\cdots,v_2^{3n}\\ \vdots\\ v_m^1,v_m^2,\cdots,v_m^{3n}\end{bmatrix} \tag{6}$$

The vector $\boldsymbol{\Phi}$ is then calculated by (7).

$$\Phi=\begin{bmatrix}\sum_{i=1}^{3n}v_1^i\\ \sum_{i=1}^{3n}v_2^i\\ \vdots\\ \sum_{i=1}^{3n}v_m^i\end{bmatrix} \tag{7}$$

The $\boldsymbol{\Phi}$ vector determines the net velocity of all markers for each frame. The total velocity of all markers is a very determinant informative value. The maximum or minimum arrays of $\boldsymbol{\Phi}$ verify the instances where the actor has reached his/her maximum angular velocity of the action. The largest column value in $\boldsymbol{\Phi}$ determines the critical feature, which is where the maximum angular velocity is observed. Assuming that during the course of an action, the markers accelerate from zero angular velocity to this value and then accelerate back to zero, time warping is carried out for each section of the action. The result is actions of the same length and temporally aligned such that sections of the action with positive angular acceleration and negative angular acceleration are aligned. Both the maximum and minimum values were tested. Figures 3 and 4 show the effect of selecting the features based on this method.

It can be noticed that in Figure 3 and Figure 4, a similar stretching as Figure 2 has occurred and the output signals are all 46 frames. Yet the alignment of the signals is regarded. The global maximum and minimum points are aligned with higher accuracy as they take place in the same time compared to Figure 2, and frames 20 to 30 have not been misaligned. Comparing Figure 3 and Figure 4, we can observe that Figure 4 shows more alignment in the case of global minimum for frames 20 to 30.

Animating the new warped signals confirms the fact that the conducted operations on the datasets have not altered the data significantly and that the primary or the secondary themes (k and r) are not altered, thus the groundwork for computing the conversion matrices has been laid successfully.

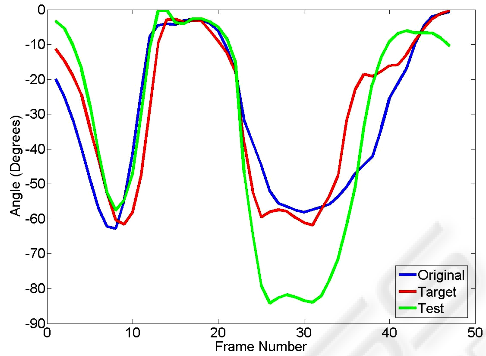

Figure 3: Signals from training, target, and test matrices after time warping using maximum velocity features.

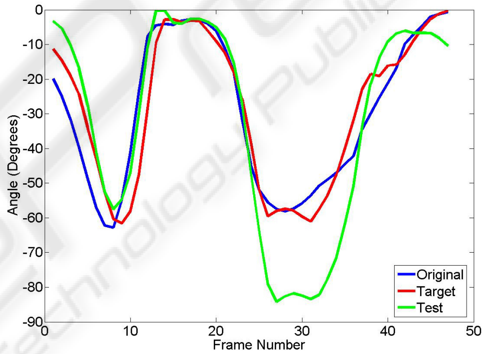

Figure 4: Signals from training, target, and test matrices after time warping using minimum velocity features.

5 CONVERSION MATRIX EXTRACTIONS AND POST-PROCESSING

With the assumption that in (3), the k and r values are unchanged and the same model is still valid, two action sequences of the same class and different secondary theme $Y_1[k,r_1]$ and $Y_2[k,r_2]$ can be modelled by (8) and (9) respectively.

$$Y_1[k,r_1]=P[k]+w[r_1]\cdot S[r_1]+e_1 \tag{8}$$

$$Y_2[k,r_2]=P[k]+w[r_2]\cdot S[r_2]+e_2 \tag{9}$$

Differentiating among Y_1 and Y_2 generates ΔY presented by (10).

$$\Delta Y=\left(w[r_2]\cdot S[r_2]-w[r_1]\cdot S[r_1]\right)+\left(e_2-e_1\right) \tag{10}$$

Assuming equal weights for the secondary themes and that only one secondary theme is present in each sequence, the differentiated secondary themes are the desired transformation matrix between the two secondary themes, *Γ[r₁,r₂]*, which is defined by (11) and computed by (12).

$$\Gamma[r_1,r_2] \equiv w[r_2]\cdot S[r_2] - w[r_1]\cdot S[r_1] \tag{11}$$

$$\Gamma[r_1,r_2] = \left(Y[k,r_2] - Y[k,r_2]\right) + \left(e_1 - e_2\right) \tag{12}$$

In simple terms, for an action sequence with only one secondary theme, the transformation matrix is derived by differentiating the base and the target sequences, along with noise elimination techniques.

It can be examined that the transformation function T defined by (13) can be applied to action sequence *Y[k,r(1:f)]* for converting the secondary theme r_i to r_j.

$$T_{r_i,r_j}[Y[k,r(1:f)]] = Y[k,r(1:f)] + w\cdot\Gamma\left(r_i,r_j\right) \tag{13}$$

To eliminate the term $(e_1 - e_2)$ from (12), low pass filtering is utilized. The affect of the noise reduction process is presented in Figure 5.

The conversion signals are a result of differentiating between two sequences of the same primary theme with different secondary themes. Due to the frame-to-frame approach for differentiation, this process is very sensitive and sharp local and global maximum and minimums are produced. Animating the final raw outcome illustrates various artefacts for different joints. Undesired motion such as tremor and rigidity in some body parts are the direct affect of the proposed techniques. Applying a low pass filter results in smoothing of the conversion data and the output signals as illustrated in Figure 5 eliminates some local maximum and minimums. Disappearance of the tremor symptoms of the animations as a result of filtering reaffirms the necessity for post-processing the data.

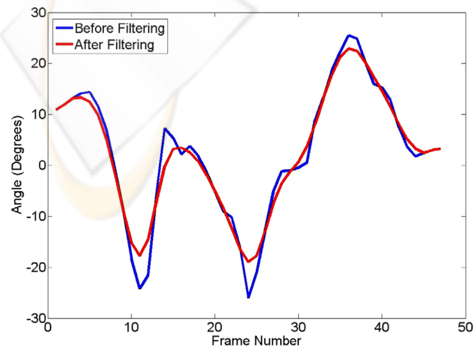

Figure 5: The effect of low pass filtering.

Another technique was also employed for transforming the secondary themes. With the assumption that one secondary theme is present in the sequence, the interpolation procedure derived by (14) and (15) will also result in *Y[k,r]* which is a sequence with the original primary action class *P[k]*.

$$Y[k,r] = \frac{\left(w_1\cdot Y_1[k,r_1] + w_2\cdot Y_2[k,r_2]\right)}{w_1 + w_2} \tag{14}$$

$$Y[k,r] = P[k] + \frac{\left(w_1\cdot w[r_1]\cdot S[r_1] + w_2\cdot w[r_2]\cdot S[r_2]\right)}{w_1 + w_2} + e \tag{15}$$

In (15) the second term represents an interpolation among the two secondary themes while the third term e is the interpolated noise signal. Also the first term *P[k]* indicates that the primary theme of the action has remained unchanged. It is very important that the secondary themes which are to be interpolated be of the same nature. For instance, they must be all related to age, or gender, or energy. Interpolating between secondary themes of different nature is meaningless as it is not logical to interpolate, for instance, between a young theme and a low energy theme. Nevertheless interpolation between a young theme and an old theme is likely to produce a mid-aged theme for the primary action class k. The same low pass filtering process is carried out to eliminate the noise from the output data.

The two techniques presented above have been implemented and the results are discussed and compared in section 6. It is important to note the fact that the conversion data do not need to include all the joint values. For instance the markers placed on the head don't have a significant impact on the secondary themes. A selected handful of markers thought to be influential to the secondary themes are altered and the rest are left unchanged, since including all the markers will only increase the run time and system error. The markers which have been selected for manipulation are on the spine, arms, and legs.

6 RESULTS AND DISCUSSIONS

As discussed in section 4 various techniques were used for determining the feature instances for the piece-wise time warping of the data. The minimum velocity-based features proved to be most suitable and are employed in the system. Successive to warping the action sequences, the transformation of the secondary themes take place using (8) to (15). The computations are carried out in Matlab and the results are transferred to Maya for visualization and further evaluation.

Interpolation between the sequences was implemented to compare with our approach. The weights of 0.5 and 0.5 for w_1 and w_2 were selected for each of the sequences to be interpolated. Figure 6 shows an original masculine jump (base) while Figure 7 presents an original feminine jumping (target) sequence successive to the time warping procedure. Figure 8 illustrates the interpolated output of the two actions. It is noticed that the action is femininized to some extent, yet creating a 100% feminine jump is only possible through setting the weight of the masculine jump to zero. Although this produces a perfect feminine jump, the dilemma is the fact that the base action (masculine) is completely excluded from the process; therefore the system is only outputting the training feminine target sequence. Figure 9 illustrates the output of the system based on computing the transformation function. In this sequence, the legs are femininized similar to the interpolation output but with few variations. The movement of the arms however, are adjusted with more significance. Overall the output is no worse, if not better, than the practical interpolation. The huge advantage remains, however, the fact that using the proposed model, we have not eliminated the base action which the secondary theme is desired to be added upon. The same affect can be seen for other primary and secondary themes and the advantage holds valid.

A rather detailed investigation for the proposed transformation technique shows that the secondary themes have been detached and added to the base samples with high precision. Figure 10 (left) shows the original masculine walk before conversion and Figure 10 (right) shows the same keyframe after the conversion. The Figure clearly shows the transformation of feminine walk to masculine walk using the model. While in feminine walk, the legs are placed sequentially in front of one another, in masculine walk, the legs are placed further apart. Also the movement of the hip is limited in masculine walk as opposed to the feminine where more movement is visible in that area. Similar manipulations are visible for other styles and other classes of action such as that illustrated in Figure 11. Figure 11 presents the conversion of low energy (tired) run (left) to a high energy run (right) using our proposed model. The movement of the legs clearly shows the successful conversion of the two styles.

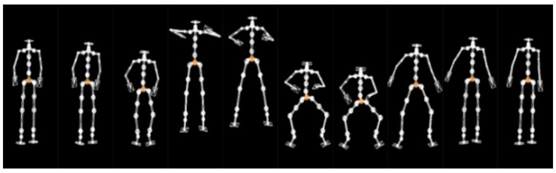

Figure 6: Original masculine jump.

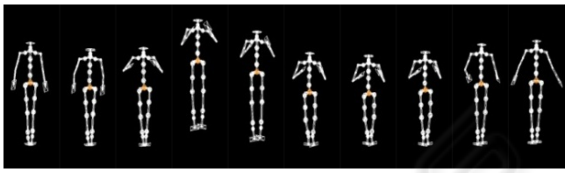

Figure 7: Original Feminine jump.

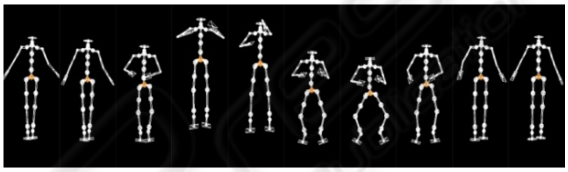

Figure 8: Interpolation output.

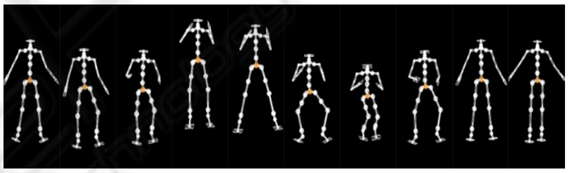

Figure 9: The output using transformation function.

To further evaluate the algorithms used to synthesize the animation, a questionnaire was formed and participants were asked to evaluate the quality of the outputted animation. The questionnaire included various enquiries regarding each primary action class and all the three secondary themes. The produced animation sequences were highly acknowledged as only a few users were unsatisfied with the results. An average 80% approval rate by the participants regarding successful style transformations using our model confirms the significance of our proposed method. Most of the concern with the animations were regarding some of the running sequences which due to the warping procedure, appeared as some sort of fast-walk, yet re-runs of the animation and further trials aided in clarifying and distinguishing between the two classes. In general, our proposed method was highly favoured over the interpolation technique.

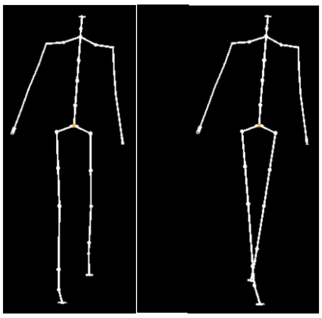

Figure 10: Original masculine walk (left), converted to feminine walk (right).

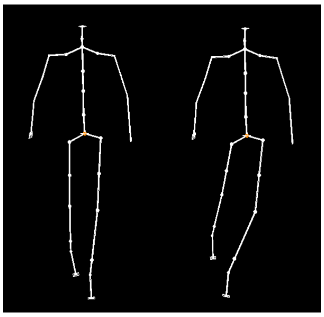

Figure 11: Original low energy run (left), converted to high energy run (right).

7 CONCLUSIONS AND FUTURE WORK

The key insights into this paper are the piece-wise time warping and features employed to conduct the warping procedure, along with the mathematical model proposed to describe human motion.

To align the motion samples correctly, critical features were computed for each action sequence. The signals were all warped such that after the warping process they all contain the same number of frames and the features be aligned i.e. they must occur at the same time. The features were computed in different fashions to search for the most functional method. Selecting the feature instance manually and using statistical analysis were the two major categories where the second proved to be much more efficient. Two different approaches were taken based on statistical feature selection: the first employed the instance when maximum net velocity is reached, and the second when the minimum net velocity occurs, where the second proved to be more effective.

The proposed mathematical model is based on dividing human motion into three components of primary theme, secondary theme, and noise. The primary theme is the main action class, for instance walking, running or jumping and the secondary theme is determined by the gender, energy, age, and other specifications of the actor. The noise is reduced by applying a low-pass filter.
Based on the model, a transformation function is defined which can extract the secondary themes from the training data and attach them to the test sequence. The computed transformation signals are used to convert any action of the same primary theme to an action containing the desired secondary theme.

For future work, further primary and secondary themes are to be tested. Actions such as idling, strafing left and right, crouching, and a collapsing or death sequence which are popular in the field of animation and game design are projected to be included. Some other secondary themes such as weight, happy/sad, and kind/angry also seem important and worth further attention.

Other transformation functions based on the proposed model could be derived and applied. As the model seems promising, further study on developing new techniques for separation of the themes is necessary.

Another area in which this research seems promising is the combination of secondary themes. It is usually necessary to produce combined styles such as old-angry, or young-feminine and etc. We are currently working on different methods for combining the transformations for each style such as weighted averaging, MIN/MAX, and rule-based operators.

REFERENCES

Brand, M., Hertzmann, A., 2000. Style Machines. In *Proceedings of the 27th International Annual Conference on Computer Graphics and Interactive Techniques*. July 2000, pp. 183-192.

Egges, A., Molet, T., Magnenat-Thalmann, N., 2004. Personalised Real-time Idle Motion Synthesis. In *Proceedings of the 12th Pacific Conference on Computer Graphics and Applications*. Oct 2004, pp. 121-130.

Fang, A. C., Pollard, N. S., 2003. Efficient Synthesis of Physically Valid Human Motion. *ACM Transactions on Graphics*. 22(3), pp. 417-426.

Hsu, E., da Silva, M., Popovic, J., 2007. Guided time warping for motion editing. In *Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation*, August 2007, pp. 45-52.

Hsu, E., Pulli, K., Popovic, J., 2005. Style Translation for Human Motion. In *ACM SIGGRAPH '05*. July 2005, pp. 1082-1089.

Hutchinson, A., 1996. *Labanotation*, Dance Books.

Kang, H., Kim, M., 2007. A Study of Walking Motion for Game Character with a Player's Emotional Factors Applied. In *5th ACIS International Conference on Software Engineering Research, Management & Applications*, Aug 2007, pp. 911-916.

Li, Y., Wang, T., Shum, H. Y., 2002. Motion Texture: A Two-Level Statistical Model for Character Motion Synthesis. *ACM Transactions on Graphics*. 21(3), pp. 465-472.

Liu, C. K., Popovic, Z., 2002. Synthesis of Complex Dynamic Character Motion from Simple Animations. In *Proceedings of the 29th International Annual Conference on Computer Graphics and Interactive Techniques*. July 2002, pp. 408-416.

Pullen, K., Bregler, C., 2002. Motion Capture Assisted Animation: Texturing and Synthesis. *ACM Transactions on Graphics*. 21(3), pp. 501-508.

Rose, C., Cohen, M. F., Bodenheimer, B., 1998. Verbs and Adverbs: Multidimensional Motion Interpolation. *IEEE Computer Graphics and Applications*. 18(5), pp. 32-40.

Safonova, A., Hodgins, J. K., Pollard, N. S., 2004. Synthesizing Physically Realistic Human Motion in Low-Dimensional, Behavior-Specific Spaces. *ACM Transactions on Graphics*. 23(3), pp. 514-521.

Sok, K., Kim, M., Lee, J., 2007. Simulating Biped Behaviors from Human Motion Data. *ACM Transactions on Graphics*. 26(3), Article 107.

Su, W., Pham, B., Wardhani, A., 2005. High-level Control Posture of Story Characters Based on Personality and Emotion. In *Proceedings of the 2nd Australasian Conference on Interactive Entertainment*, Nov 2005, pp. 179-186.

Tanco, L. M., Hilton, A., 2000. Realistic synthesis of novel human movements from a database of motion capture examples. In *Proceedings of Workshop on Human Motion*, July 2000, pp. 137-142.