

Temporal Transfer of Locomotion Style

Yejin Kim, Myunggyu Kim, and Michael Neff

Timing plays a key role in expressing the qualitative aspects of a character's motion; that is, conveying emotional state, personality, and character role, all potentially without changing spatial positions. Temporal editing of locomotion style is particularly difficult for a novice animator since observers are not well attuned to the sense of weight and energy displayed through motion timing; and the interface for adjusting timing is far less intuitive to use than that for adjusting pose. In this paper, we propose an editing system that effectively captures the timing variations in an example locomotion set and utilizes them for style transfer from one motion to another via both global and upper-body timing transfers. The global timing transfer focuses on matching the input motion to the body speed of the selected example motion, while the upper-body timing transfer propagates the sense of movement flow — succession — through the torso and arms. Our transfer process is based on key times detected from the example set and transferring the relative changes of angle rotation in the upper body joints from a timing source to an input target motion. We demonstrate that our approach is practical in an interactive application such that a set of short locomotion cycles can be applied to generate a longer sequence with continuously varied timings.

Keywords: Timing, temporal transfer, locomotion style, character motion, succession.

Manuscript received Jan. 29, 2014; Dec. 27, 2014; accepted Jan. 2, 2015.

This work was supported by the Ministry of Science, ICT and Future Planning (MSIP) and Institute for Information and Communications Technology Promotion (IITP) through Digital Contents Research and Development Program 2013 (10044309, Development of Performance Analysis and Generation Technology for Experiencing, Learning and Creating Traditional and Popular Dances).

Yejin Kim (corresponding author, yejink@etri.re.kr) and Myunggyu Kim (mgkim@etri.re.kr) are with the SW-Content Research Laboratory, ETRI, Daejeon, Rep. of Korea.

Michael Neff (mpneff@ucdavis.edu) is with the Department of Computer Science, University of California, Davis, USA.

I. Introduction

From an idle stance to a graceful ballet, there are a wide variety of human motions captured for different synthesis purposes. Among them, human locomotion is one of the most widely used activities in interactive applications such as computer games and virtual agents, yet it often requires laborious effort and expertise in motion properties to transform input data into a desired style [1]. This is especially true for a novice animator, whereby such a person can be faced with editing a 3D articulated character consisting of a high number of degrees of freedom (DOF) and needing to maintain correlations between body parts throughout the locomotion sequence.

Recent data-driven approaches [2]–[6] focus on how to satisfy the pose constraints imposed by an animator and generate stylistic variations from the example motion set. These approaches are particularly focused on the timing adjustments to make a quick style change for an input motion. As noted in [7], observers feel very differently about the sense of weight and energy shown by a character that moves through the same spatial positions when there are temporal changes. Furthermore, the qualitative aspects of style, such as the emotional state and particular role of a character, are influenced arguably more by timing variations than by postural configurations [8]. For example, the number of in-between frames between two keys in a walk can specify a degree of tiredness or depression throughout the motion sequence. Unfortunately, current approaches pay less attention to reusing such stylistic examples in expressive locomotion generation. They require users to set a number of key frames or control handles to impose pose constraints, which can be a tedious job for a large example set with cycled motions like human locomotion. Moreover, their temporal edits mainly target controlling the variable speed of the motion in the output

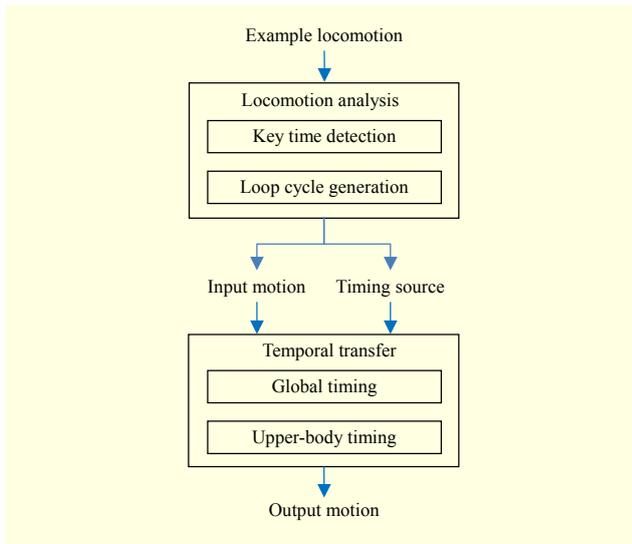


Fig. 1. Overview of temporal transfer of locomotion style.

sequence and neither effectively capture the timing variations within the example sequence itself nor utilize such timing differences for the style transfer.

In this paper, we propose an editing system for locomotion style that provides temporal adjustments to an input motion via global and upper-body timing transfers from an example set. As shown in Fig. 1, our system first constructs loop cycles from the example set and uses them as an *input target* and *timing sources* — an input target is the motion whose timing we want to transform with the style edit, while a timing source is the motion that we want to transfer the timing from. Thus, the output motion has spatial qualities of the input and timing features of the example source.

During global timing transfer, we extract the timing variations in the example set based on the duration between the detected key times and apply the extracted timing distribution to the input motion. This matches the input motion to the variable body speed of the selected example motion and will contain the overall sense of weight and energy observed in the example. In addition, through an upper-body timing transfer, the system propagates the sense of movement flow throughout the upper body, which is often referred to as *succession* [9]–[10]. Functionally, succession is the process of movement passing from one joint to the next rather than all movement starting and stopping at the same time. We transfer this succession by capturing the relative changes of angle rotation in the upper body joints from the timing source and then applying them to the input motion with a scaled amount of joint rotation.

Our system makes two main contributions. First, we extract the timing variations from the example set and apply them to a motion cycle for style modification without destroying spatial details and constraints present in the original motion. This makes

our approach practical for interactive applications such that a set of short locomotion cycles can be applied to generate a longer sequence with continuously varied timings. Second, to generate a convincing output, we provide a simple but effective timing transfer processes for the variable body speed and succession in the upper body part. During these transfers, the system requires minimum user intervention and provides quick style changes. Our experimental results show that an animator can generate stylistic variations from the example set simply by selecting the input and timing sources that exhibit the desired style.

The remainder of this paper is organized as follows. We begin with a survey of previous approaches for temporal edits on motion style in Section II. The locomotion analysis is explained in Section III, and the temporal transfers are detailed in Section IV. After experimental results are demonstrated in Section V, we conclude this paper with a discussion of potential improvements in Sections VI and VII.

II. Related Work

Over the years, a variety of research has been conducted that uses timing aspects of motion as an editing tool for changing the style in the motion synthesis. Witkin and Popovic [11] used spline-based interpolation of key frames, which modifies an example motion to achieve a certain pose at a given time point while preserving original content of the motion elsewhere. A sketch-based system was adopted by Terra and Metoyer [12] to set object timings on the motion path directly from a user's input device. Incorporating the timing changes of example motions into the physics-based system, McCann and others [6] focused on maintaining physical validity of the temporally modified motion by using the optimization of the objective function; a function that balances the timing and torque controls. In an attempt to speed up the timing modification process, Hsu and others [3] compressed or expanded an input motion temporally to follow the local timings in the example motion. For this, they used an objective function to access the similarity between two motions and performed a constrained path search on the discretized frames. On the other hand, Coleman and others [2] provided a single-pose representation, which is designed to control the timing variations of each joint over multiple frames concurrently. Based on a multivariate statistical method, Lau and others [4] trained a generative model with a relatively small number of examples, which synthesizes temporal and spatial variants of the input motions. Recently, Lockwood and Singh [5] proposed a motion path-based editing system that first modifies spatial data of a motion via a manipulation of the control points on the path and then adjusts the timings of footsteps to maintain temporal similarity with the original motion. Most of these approaches, however,

have paid little attention to capturing and applying the temporal variations from example motions for style transfer.

Within a procedural animation system, Chi and others [13] provided explicit parameters such as frame duration and a velocity function to show different timing patterns of arm movements. In a similar sense, Neff and Fiume [14] introduced a kinematic approach to improve the sense of flow of gesture movements by delaying joint rotations to generate succession in transitions between key poses.

Our research is also related to the techniques that replace specific body parts or the DOF of one motion with another to expand the size of a motion database. Pullen and Bregler [15] added high-frequency details of an example motion to the output by filling the sparse DOF specified in a key-framed manner. Al-Ghreimil and Hahn [16] extracted a partial motion of an upper body performing a throwing motion and combined it with walking data. Ikemoto and Forsyth [17] designed a rule-based classifier to transplant limbs in an example set. Ashraf and Wong [18] divided a character body into an upper and a lower half and maintained consistently synchronized locomotion between the two halves via decoupled interpolation. Similarly, Heck and others [19] introduced a layered approach that focuses on preserving the cross-body correlation for the upper and lower body combination. Later, Oshita [20] tried to maintain the correlation by adding torso vibrations, extracted from a given action, to the lower body motion. In our approach, we do not replace DOF between the motions to generate the variants of the input motion; instead, we focus on modifying the input style via the timing transfers without worrying about spatial data changes.

III. Locomotion Analysis

Human locomotion is a complicated but constrained activity. Physically, one walks by propelling the lower limbs in a cyclic pattern to maintain balance above the ground. This makes a locomotion cycle, called a *loop cycle* in our work, a natural choice as a compact and structured representation for style editing since it can be concatenated repeatedly and will yield smooth locomotion without further blending during the locomotion synthesis. Assuming each example clip starts and ends with the same foot step, a stance period can be defined by detecting the important moments of foot contact on the ground. We call such moments *key times* and use them to set correspondences between the examples for the timing transfer process.

1. Key Time Detection

It is a laborious task to tag every key time manually, even for

Table 1. Detection of key times: positional acceleration of ankle and toe joints crossing zeroes or being near zero is used to detect key times for each foot stance.

Key times	Support foot acceleration	Swing foot height
	Ankle, toe	Ankle
Start stance (SS)	$\approx 0, > 0$ or $> 0, \approx 0$	—
Mid-stance (MS)	$\approx 0, \approx 0$	Highest
End stance (ES)	$> 0, > 0$	—

a short locomotion sequence. Furthermore, the precise starting moment for each stance is ambiguous to detect with a fixed threshold value as the example motion contains noise and retargeting errors [21]. For this reason, we first attempt to reduce the noise for the foot joints with a low-pass filter, the Savitzky–Golay filter [22], which smooths out local peaks of noise from motion curves of foot joints while maintaining the high peaks that possibly belong to a frame not belonging to a stance period. For our example set, a filtering window size consisting of ten frames and a smoothing polynomial of degree four worked well for preserving the global shape of significant peaks in target data [23].

As shown in Table 1, we detect three different key times to define a period for each foot stance: SS, MS, and ES. Here, SS is the start of the stance period; that is, the point where the leading part (for example, heel or toe) of the foot initially contacts the ground. MS is the moment when a swing foot reaches its highest point while the support foot rests on the ground. ES is the end of the stance period; that is, the moment when contact is broken between the foot and the ground. To effectively capture the key times for the stance period, we applied zero-crossings of acceleration of two joints, the ankle and toe, from the support foot and the height position from the swing foot, as shown in Table 1. For the zero-crossings of joints, a spatial filter, Laplacian of Gaussian, is applied to detect the sign change in the acceleration data. It is noteworthy that there are two different cases for SS detection due to the heel-strike or toe-strike moment of the foot. When there are multiple detections for each key time, we select the one that has the lowest height position or let an animator select one.

2. Loop Cycle Generation

Given the key times, a loop cycle can be constructed by blending the first stance with the second stance, as shown in Fig. 2. Based on the DTW technique [24], we set the correspondence between two stances with the blending window size, $\min(ES_i - SS_i, ES_{i+1} - SS_{i+1})$. Each of the in-

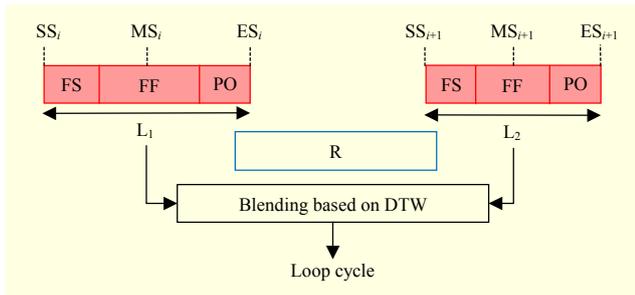


Fig. 2. Specification of a loop cycle: two foot stances, L_1 and L_2 , are dynamically time-warped (DTW) and blended to replace the first stance (L_1), which creates the smooth loop continuity.

between frames between two key times is generated by linearly interpolating their root positions and performing spherical linear interpolation [25] on the joint rotations.

IV. Temporal Transfer

Table 2 shows the timing variations in the example set in terms of time spent (for example, in-between frames used) by different foot phases: foot strike (FS), foot flat (FF), and push off (PO), as shown in Fig. 2. Similar to the key times detected for a foot stance in Section III-1, the aforementioned foot phases represent states comprising foot joints that are in contact with the ground, within each foot stance: FS is a phase starting at SS and going to FF, where FF starts with two foot joints (for example, toe and ankle) having zero-crossings. After MS, PO starts with one of the foot joints being pushed off from the ground until ES. The stylistic differences in the example set can be observed from these foot phases. For example, the “sneaky” motion spends considerably more time on the FS and PO phases than the “tired” one does due to its cautious movements taking longer on the toe or heel areas. It is from this perspective that we match the variable body speed of the timing source to an input motion based on the duration ratios of corresponding phases between two motions. In our approach, we use the phase durations between SS, MS, and ES, instead of FS, FF, and PO, since they are better suited for controlling the durations of leg rising and lowering the output motion.

1. Global Timing Transfer

Given the shape of the motion path in the space, the distance-time function, which relates time to distance traveled, can be defined to set the body speed along the curved path. This distance along a curve is defined as *arc length*, and it is specified as a function of time to parameterize the curve within a range of a parameter value, t , where $t \in [0, 1]$. However, for an arbitrary path, this relationship between the distance along

Table 2. Number of in-between frames used by each foot stance in the example set.

Examples	Left foot			Right foot		
	FS	FF	PO	FS	FF	PO
Basic	28	18	37	24	16	38
Female	42	14	49	38	12	42
Jogging	18	19	16	17	19	15
Bouncy	10	23	25	12	24	28
Confident	14	23	25	12	20	33
Sneaky	71	20	58	68	14	52
Energetic	13	29	20	10	28	15
Tired	20	69	13	20	62	15

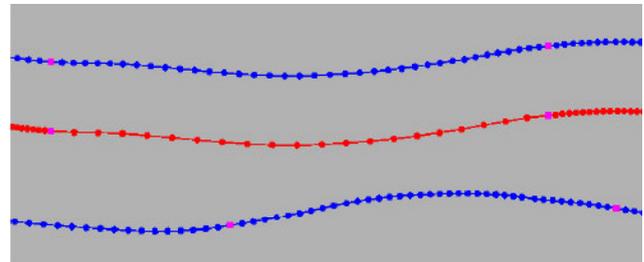


Fig. 3. Temporal spacings between in-between frames (sphere dots) based on the same parametric values used for the arc length traveled between two example motions. Each curve shows a motion path traveled by a root joint. Here, the top and bottom paths (blue) are drawn from two types of walk, basic and sneaky, respectively. Middle path (red) is generated by locating the corresponding positions on the basic walk by using the parametric values from the sneaky walk. Square dots (pink) indicate the corresponding key times between the motions.

the curve and t does not always result in the desired body speed. For example, when the path curve is defined for the root joint, we can locate the corresponding positions on the curve from one motion to another by using the same t values and interpolation of two neighboring points on the curve (for example, arc length-based approach). However, as seen in Fig. 3, using this t value from a motion with higher body speed for a motion with lower speed, the duration between two consecutive frames become longer, and a viewer can perceive it as a discontinuity in the motion. The distribution of phase durations varies in stylistically different motions, as shown in Fig. 4. This means that the portion of arc length covered in a phase will be different across these motions. Applying the timing from a phase in one motion to the same phase in another motion can lead to unnaturally slow or fast velocities. Thus, the direct transfer of the arc lengths between two motions only works reasonably when there is a similar distribution of phase

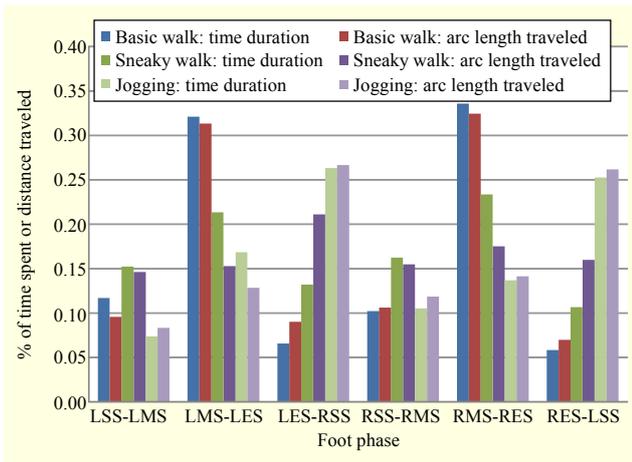


Fig. 4. This shows the distributions of various foot phases in terms of the percentage of time spent or distance travelled by the root joint for various examples. Here, SS, MS, and ES denote the key times for start stance, mid-stance, and end stance in a left- (L) or right- (R)-foot stance, respectively.

durations for the motions. However, this is not a typical case as the stylistic motions show large velocity changes throughout the motion sequence. For this reason, we match the ratio of velocities for each phase (for example, between two key times) in the output to be the same as the ratio of velocities in the timing source; thus, the duration of the output should be the same as the duration of the timing source.

Formally, let e_i and o_i be the number of in-between frames and the distance traveled along the arc length of the i th phase in the input motion, respectively, where $i \in [1, \dots, N_p]$ and N_p is the total number of phases. Similarly, the corresponding j th phase in the timing source defines \tilde{e}_j and \tilde{o}_j . Since we want to adjust the ratio of the input motion, e_i/o_i , to be the same as that used for the timing source, we redistribute the total time spent by the timing source to each phase duration in the input motion based on its new ratio, \tilde{e}_i , as follows,

$$\hat{e}_i = \frac{\tilde{e}_i}{\sum_1^{N_i} \tilde{e}_i} N_T \quad \text{for} \quad \tilde{e}_i = o_i \frac{\tilde{e}_i}{\tilde{o}_i} N_T. \quad (1)$$

Here, N_i and N_T are the number of total in-between frames in the input and timing source, respectively. It is noteworthy that \tilde{e}_i is the number of frames required to move at the average velocity of the timing source over the distance for each phase in the input, while \hat{e}_i ensures that the output duration is the same as the timing source.

Given \hat{e}_i for the i th phase in the output motion, we now need to redistribute the in-between frames. From the parameterized curve, we can approximate a corresponding arc length, \hat{a}_o , from \hat{u}_k , by computing the linear distance between two neighboring points in the input motion as follows:

$$\hat{a}_o = a_k + \frac{\hat{u}_k - u_k}{u_{k+1} - u_k} (a_{k+1} - a_k), \quad (2)$$

where \hat{u}_k and u_k are normalized by \tilde{e}_i and e_i , respectively, such that $\hat{u}_k, u_k \in [0, 1]$ for each phase. Here, a_k is an arc length traveled by u_k in the corresponding phase of the input motion. The actual location on the output curve for a root joint is estimated by interpolating the two nearby frames in the input motion, located at a_k and a_{k+1} , with the weight derived from $(\hat{a}_o - a_k)/(a_{k+1} - a_k)$.

2. Upper-Body Timing Transfer

Once the body speed of the input motion is matched to the selected timing source, the timing of the upper body is transferred from the source to the input motion. In this process, we mainly focus on capturing the succession starting from the root and passing throughout the upper body parts. The timing changes for the lower body parts can be adjusted by skipping or delaying the time spent to lift or to lower a leg; however, these can potentially destroy some of the key physical characteristics and constraints embedded in the original motion [5], [19]. On the other hand, the spatial and temporal attributes of a motion in the upper body are less constrained and convey more stylistic variation for expressive locomotion [18].

As shown in Fig. 5, our timing transfer for the upper body between two motions is performed with two separate variables; that is, joint rotation shift, $\Delta\beta$, and scale, α . The rotation shift is achieved by moving the time between the peak rotation of child and parent joints in the input motion to align with the ones in the timing source. These joint rotations are further edited by scaling their range of movement in the upper body based on the relative changes in the amplitude between the input and timing source.

3. Joint Rotation Shift

In the articulated body model, the succession amount passing throughout the upper body is defined by the relative offset in the start and stop times (or peak values) of rotations between the child and its parent joints [13]–[14]. Inspired by this, the temporal difference between the two peak points on the child and its parent joint curves are compared between the two motions to determine the joint rotation shift in our system. As shown in Fig. 5, $\Delta\beta_i$ and $\Delta\beta_T$ are the timing differences between those joints for the input and timing source, respectively. We first shift the child curve by $\Delta\beta_i$ to align its peak with the parent curve; or vice versa. When N_i and N_T are the total number of in-between frames in the input and timing source, respectively, the temporal order and shifting amount, $\Delta\beta_i'$, for the child curve for the input motion is determined by

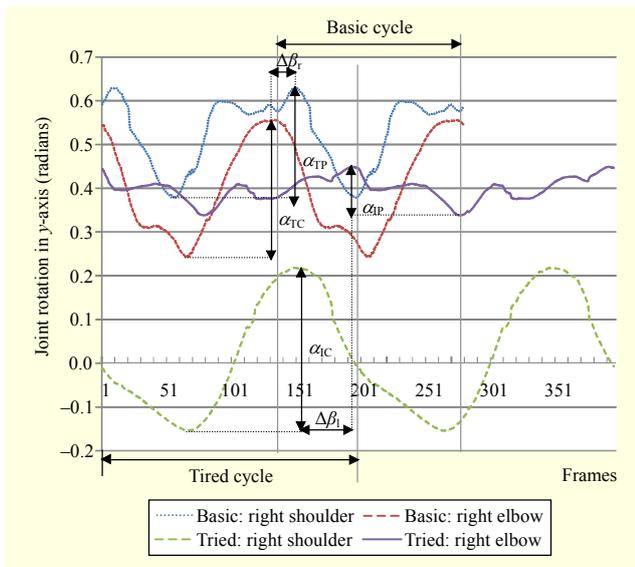


Fig. 5. Joint rotation shift and scale: shoulder and elbow joints in the two cycles of a loop sequence are plotted from the basic and tired walks, respectively. Here, shifting amount, $\Delta\beta$, is defined as the distance between two extreme points of the shoulder and elbow joints, while peak-to-peak amplitude, α , will be used for the scaling amount.

$$\Delta\beta'_I = N_I \frac{\Delta\beta_T}{N_T}, \quad (3)$$

where the shift direction (for example, temporal order) is determined by the sign of $\Delta\beta_T$ in the timing source. For example, in the case of a negative sign, we shift the child curve to the left to place its peak in front of the parent curve. In the case of multiple peaks in the cycles, we shift the rotation curves based on its first detected peak. In our experiments, we used the highest peak points for the input and timing sources since it is a moment when the joint rotation starts to change its direction (for example, the succession starts to propagate throughout the upper body); however, the lowest peaks also work for the same shift. Since a loop cycle is used for the timing transfer; hence, shifting does not disturb the smooth continuity in the input motion. For multiple cycles, the same shift is applied for the entire motion sequence, as shown in Fig. 6.

While the joint rotation shift is directly related to the succession transfer throughout the upper body parts, it does not adjust the spatial movements of the input motion in correspondence with the physical characteristics shown from the timing source. For example, when the basic motion is transferred to the tired motion, the range of the arm swings should be reduced for the tired state. To maintain a plausible range of movements in the output motion, we scale the joint rotations in the upper body based on the peak-to-peak amplitude between the input and timing source, as shown in

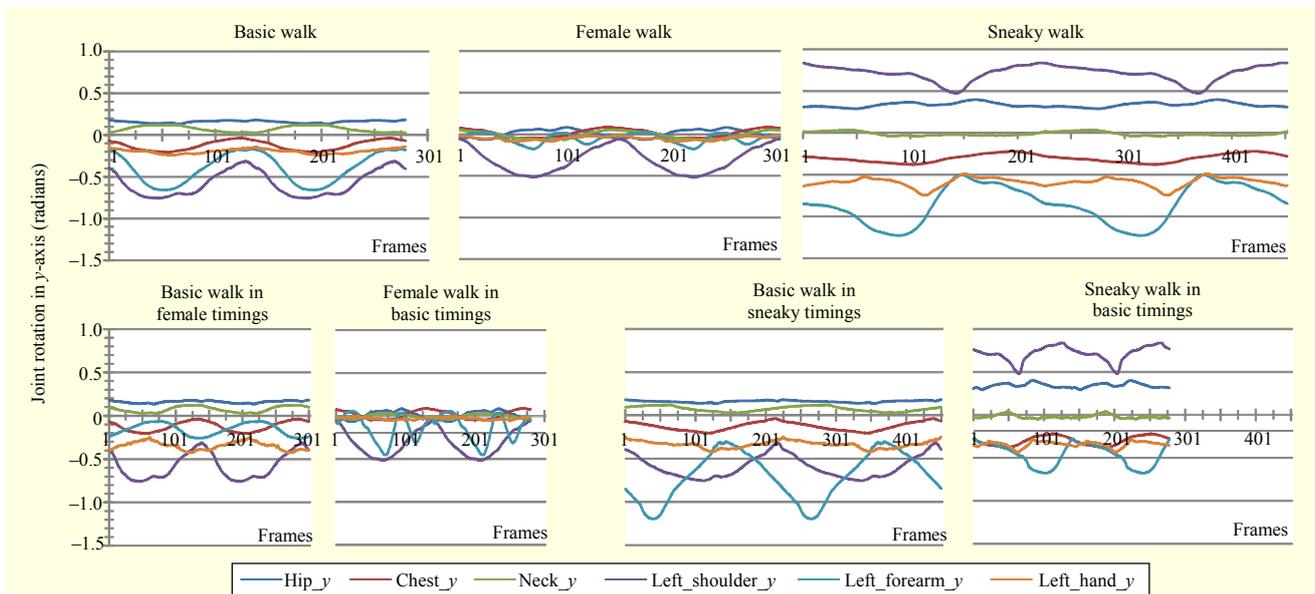


Fig. 6. Comparison of temporal transfers between the example motions: in the basic walk, peak rotations for the upper body, especially the arm joints, happen at similar times. In the basic walk in female timings, they are now shifted to the left side and their amplitudes are scaled in correspondence with the pattern observed in the female walk. In the female walk in basic timings, peak rotations are shifted to the right side, while some of the amplitudes are scaled up largely (that is, the forearm joint) for bigger arm swings observed in the basic walk. In the basic walk in sneaky timings, peak rotations and their amplitudes are shifted and scaled in correspondence with the pattern observed in the sneaky walk. For example, peak rotations for the forearm joint now happen earlier than its parent, the shoulder joint, while their amplitudes are scaled up relatively. This happens in the opposite way with the sneaky walk in the basic timings. In the graphs, selected joint rotations are traced for the first two cycles.

Fig. 5. This is performed in a similar way as the shifting operation; that is, between the child and its parent joint, a scaling amount, δ , for the input motion is estimated based on the ratio of the amplitudes in the rotation curve as follows:

$$\delta = \begin{pmatrix} \alpha_{IP} \\ \alpha_{IC} \end{pmatrix} \begin{pmatrix} \alpha_{TC} \\ \alpha_{TP} \end{pmatrix}, \quad (4)$$

where α_{IC} and α_{IP} are the peak-to-peak amplitudes for the child and its parent joints in the input motion, respectively, and similarly, α_{TC} and α_{TP} are defined for the timing source. Here, we set $\delta = 1$ if the amplitude for the child or its parent joint is constant. To scale the range of the movements in the upper body, we multiply δ with the joint rotation values in the input motion.

V. Experimental Results

Our experiments were performed on an Intel Core2Quad™ 2.4 GHz CPU with 8 GB memory. The skeleton structure used for all outputs consists of 50 DOF: 6 for the root translation and orientation; 12 for the torso and head orientation; 9 for each arm orientation; and 7 for each leg orientation. All example motions are captured at a rate of 120 fps. The system is best understood through examples of its use, as described below and included in the accompanying video.

1. Temporal Transfer

Our experiments start with comparing two outputs generated from different global transferring methods. For this, we used the basic walk as the input motion and female walk as the timing source to show their differences in the temporal transition between the key times. Applying the arc length-based method, the output shows discontinuities between the foot stances due to the sudden changes in the speed and linear spacings between the in-between frames, as shown in Fig. 7. On the other hand, the ratio-based method provides gradual

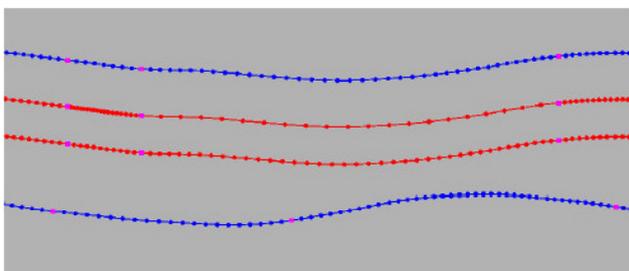


Fig. 7. This shows distributions between the time spent and arc length traveled by the root joint from various examples. Here, SS, MS, and ES denote the key times for start stance, mid-stance, and end stance, in a left (L) or a right (R) foot stance, respectively.

spacings and varies the speed smoothly from one stance to another, which is similar to the pattern observed in the captured motion. This justifies the use of the ratio-based method over the arc length-based method.

Next, we show the style transfers between two example motions. For these edits, we selected one motion as an input target motion and another as a timing source from the example set. The variable body speed and upper-body timings of the source are transferred into the input motion. At first, we used the basic walk as input and the female walk as the source. The output motion is shown in Fig. 8; the basic walk in the timings of the female walk shows some of the temporal characteristics shown in the female walk such as the modulation in the torso movements and relaxed arm swings with reduced range compared to the basic walk. When we exchanged the input and timing source, the output (the female motion in the timings of the basic walk) showed an increasing rigidness throughout the torso and forearm bends, which is reminiscent of the temporal features observed in the basic walk. Figure 6 compares the joint rotations between these motions. Next, we performed similar transfers between the basic and sneaky motions. Even though there is a large difference in the body speed for each phase between these two motions, the output motions exhibit stylistically very different movements from their input motions. As seen in Fig. 8, we can observe the sense of caution and hesitation, not present in the original basic walk, from the basic walk in the sneaky timings. On the other hand, there is rush and

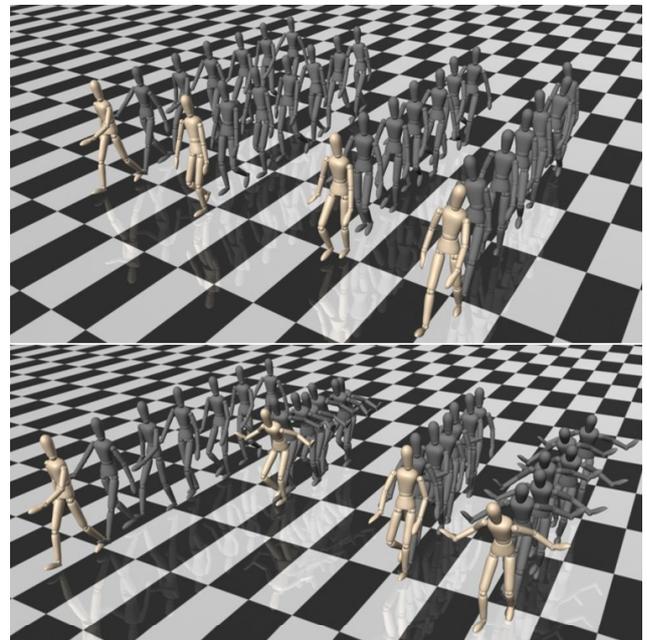


Fig. 8. Output comparisons between the basic and female timings (top) and between the basic and sneaky timings (bottom). Two example motions and their outputs are shown from left to right, respectively.

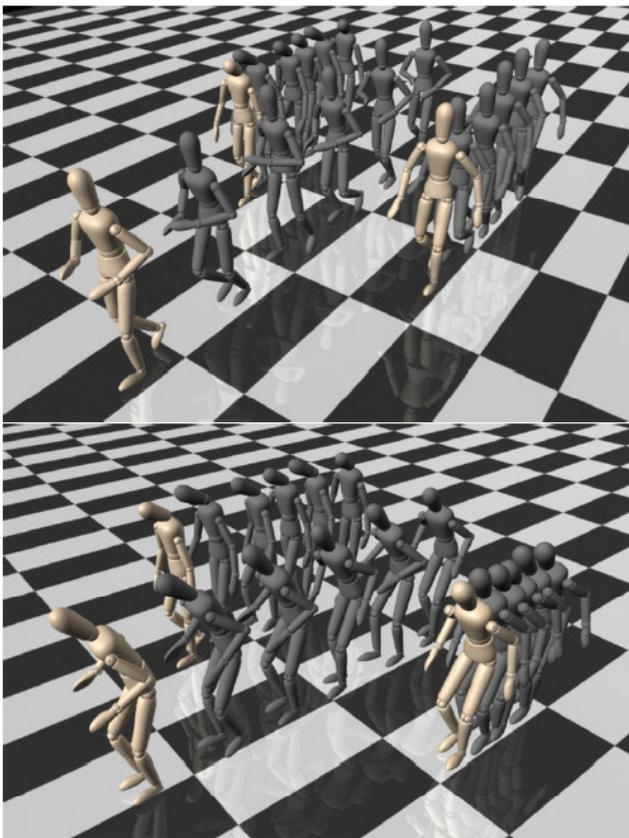


Fig. 9. Outputs in tiredness: jogging (top) and nervous (bottom) motions are edited with the tired timings, respectively. Here, timing source, input, and output motions are shown from left to right for each output.

haste in the sneaky walk with the basic timings. This output is further edited to generate a slightly pulled-back version by slowing down the step-down periods of the feet. For this, we simply extended the duration between MS and ES by 1.5 times and the duration between ES and SS by 2 times for each foot. Figure 6 compares the joint rotations between these motions.

In addition, various style transfers are performed from the example set; that is, the jogging and nervous motions in the tired timings, the balancing motion in the energetic timings, and the ape-like motion in the balancing timings. As seen in Figs. 9 and 10, all of these outputs show plausible motions that are similar to the input motion spatially and to the timing source motion temporally. For example, the physical state, such as tiredness (from the tired motion) or haste (from the energetic motion), not present in the input motion, is introduced into the output motion only by editing the temporal attributes of a timing source. However, as seen in Fig. 10, the ape-like motion in the balancing timings shows that the output lacks balancing movements; rather, it looks more like a tired ape-like motion. This is mainly because the lower-body movement in this source motion contains key characteristic features of the

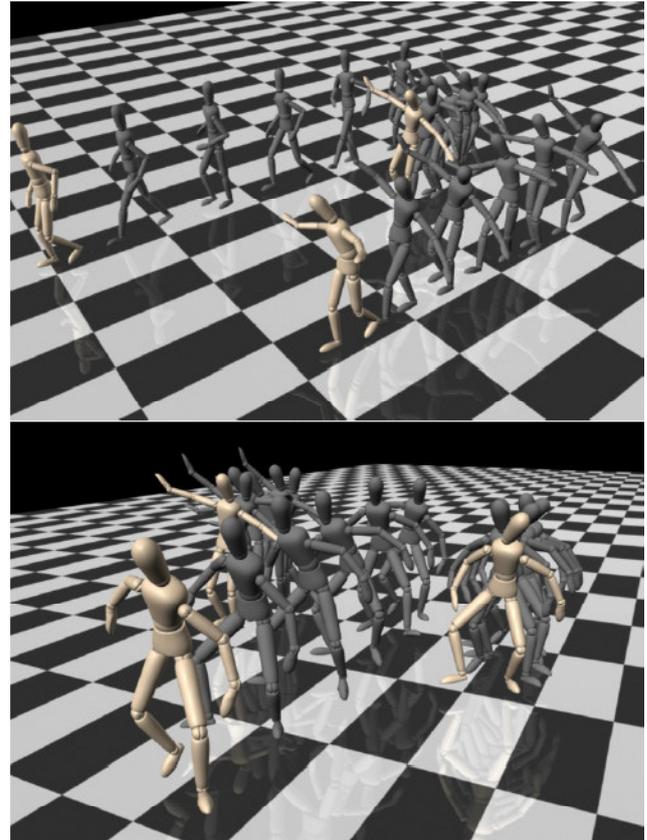


Fig. 10. Outputs with balancing motions: balancing motion in the energetic timings (top) and ape-like motion in the balancing timings (bottom) are shown, respectively. Here, timing source, input, and output motions are shown from left to right for each output.

balancing action. Nevertheless, our temporal transfer between different examples proves to be very useful to generate output that exhibits the style characteristics of the selected example without destroying the spatial details and constraints of the original motion.

2. Locomotion Synthesis on Arbitrary Path

We demonstrate the usefulness of our style transfer in the interactive applications by applying a set of short loop cycles to generate a longer sequence with continuously varied timings. For this, we selected the basic walk as an input motion and a set of timing sources from basic, sneaky, jogging, tired, and energetic walk. At first, five loop cycles are generated by applying temporal transfers, such as the basic walk in the energetic timings, and other timing sources. To generate an output of a longer sequence, we applied the locomotion synthesis on an arbitrary motion path suggested by [24], which concatenates a series of the loop cycles on the path specified by an animator. As shown in Fig. 11, the character follows the given path while showing a series of different emotional and

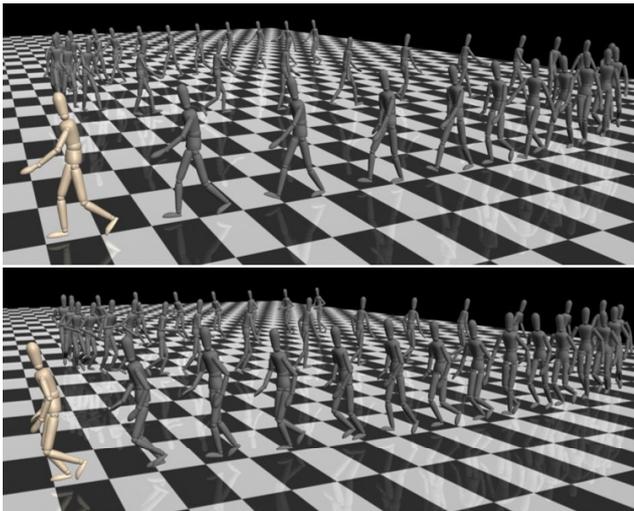


Fig. 11. Locomotion synthesis on arbitrary path: basic (top) and jogging (bottom) sequences follow a given path with a set of timing sources.

Table 3. Perception test for a set of output styles generated from the following timing sources: energetic (E), tired (T), sneaky (S), nervous (N), ape-like (A), and balancing (B) motions.

Timing sources	Perception (%)					
	E	T	S	N	A	B
Energetic	84	0	12	4	0	0
Tired	0	80	4	12	0	4
Sneaky	0	20	76	4	0	0
Nervous	0	16	0	84	0	0
Ape-like	16	0	16	8	60	0
Balancing	28	0	8	12	8	44

physical states in the output sequence, derived from the timing sources. Similarly, an additional output is generated by using the jogging motion as input and a set of timing sources from the jogging, sneaky, ape-like, tired, and female walks.

3. User Case Study

Twenty-five college students were recruited for the perception task to test the effects of our system. We used a set of motion clips that contains a basic locomotion with different timing sources from the example set (see Table 3). Each subject was asked to select the most likely style perceived after watching a clip several times. Table 3 shows the selection results in terms of percentages; that is, the number of votes received for each style is divided by the total number of subjects. It is noteworthy that for each of the timing sources, a significant majority of the twenty-five subjects correctly

perceived the style of the judged motion to be the same as that of the timing source used in its creation, except the balancing motion, where spatial data is the dominant factor to decide this particular style.

VI. Conclusion

Style is probably the most important element in expressive character animation as it conveys to the audience the distinctive information about a character and makes animation content meaningful. Nevertheless, editing a wide variety of styles from a given motion via a number of low-level control parameters or DOF is a laborious and time-consuming task for most users who are not familiar with human motion properties and lack years of animation training. This is especially true for editing the temporal aspect of motions, as noted by [7] with “The emotional state of a character can also be defined more by its movement than by its appearance, and the varying speed of those movements indicates whether the character is lethargic, excited, nervous or relaxed.”

In this paper, we have presented a temporal editing system that captures timing variations from an example set and applies them to an input motion for style transfer. For this, we provided a simple but effective method to construct a loop cycle for each example motion and a timing transfer process for the variable body speed and succession in the upper body from one motion to another. During these transfers, the system requires minimum user intervention and provides quick style changes. A user study confirms that our system is able to effectively transform an input motion into a desired style.

In conclusion, the proposed system is applicable for interactive systems, where a character locomotion is typically directed to follow a motion path, as demonstrated in our results. In addition, we expect that our system can contribute to expanding the range of motion data available in a quick, cost-effective way by editing the existing data into a desired style without destroying physical details in the original motion.

VII. Discussion

Some of the limitations in our system are discussed in this section. The proposed system focuses on expressive locomotion generation based on specific example data, such as repetitive motions like walking. This immediately imposes a limitation on its usage on other types of motion data. In cyclic motions like walking or jogging, the correlations between the body parts are generally strong throughout the motion sequence, especially the arm and leg swings. This will not be the case for acyclic motions such as ballet and dance. While specific motions may have strong correlations, their

correlations may vary from motion to motion; thus, the motions certainly will not be cycled in the same way as cyclic ones. Using such types of motions may require a more sophisticated strategy [26] to detect important moments of style, not only from the foot-plants, but also from other body parts, such as the head and hands, to establish more comprehensive correlations throughout the body.

Our temporal transfers can generate plausible output motion from any example set; however, its style might not be one that an animator intended. This is especially true when the lower-body movements in the timing source contain key characteristic features of the desired style. For example, the ape-like motion generated with the balancing timings, as shown in Fig. 10, appears more like a tired motion due to the missing spatial data from the source — the balancing action. Thus, our approach works best for motions that convey stylistic activities mainly from the upper body parts.

References

- [1] O. Johnston and F. Thomas, “*Disney Animation: The Illusion of Life*,” New York, NY, USA: Abbeville Press, 1981.
- [2] P. Coleman et al., “Staggered Poses: A Character Motion Representation for Detail-Preserving Editing of Pose and Coordinated Timing,” *ACM SIGGRAPH/Eurograph. Symp. Comput. Animation*, Dublin, Ireland, 2008, pp. 137–146.
- [3] E. Hsu, M.D. Silva, and J. Popovic, “Guided Time Warping for Motion Editing,” *ACM SIGGRAPH/Eurograph. Symp. Comput. Animation*, San Diego, CA, USA, 2007, pp. 45–52.
- [4] M. Lau, Z. Bar-Joseph, and J. Kuffner, “Modeling Spatial and Temporal Variation in Motion Data,” *ACM SIGGRAPH*, vol. 28, no. 5, 2009, pp. 1–10.
- [5] N. Lockwood and K. Singh, “Biomechanically-Inspired Motion Path Editing,” *ACM SIGGRAPH/Eurograph. Symp. Comput. Animation*, Vancouver, Canada, 2011, pp. 267–276.
- [6] J. McCann, N.S. Pollard, and S. Srinivasa, “Physics-Based Motion Retiming,” *ACM SIGGRAPH/Eurograph. Symp. Comput. Animation*, Vienna, Austria, 2006, pp. 205–214.
- [7] J. Lasseter, “Principles of Traditional Animation Applied to 3D Computer Animation,” *ACM SIGGRAPH Comput. Graph.*, Anaheim, CA, USA, vol. 21, no. 4, July 1987, pp. 35–44.
- [8] H. Whitaker and J. Halas, “*Timing for Animation*,” London, UK: Focal Press, 1981.
- [9] R.V. Laban and L. Ullmann, “*The Mastery of Movement*,” London, UK: Northcote House, 1988.
- [10] T. Shawn, “*Every Little Movement: A Boot about Delsarte*,” Hightstown, NJ, USA: Princeton Book Co., 1976.
- [11] A. Witkin and Z. Popovic, “Motion Warping,” *ACM SIGGRAPH*, Los Angeles, CA, USA, 1995, pp. 105–108.
- [12] S.C.L. Terra and R.A. Metoyer, “Performance Timing for Keyframe Animation,” *ACM SIGGRAPH/Eurograph. Symp. Comput. Animation*, Grenoble, France, 2004, pp. 253–258.
- [13] D.M. Chi et al., “The EMOTE Model for Effort and Shape,” *ACM SIGGRAPH*, New Orleans, LA, USA, 2000, pp. 173–182.
- [14] M. Neff and E. Fiume, “Aesthetic Edits for Character Animation,” *ACM SIGGRAPH/Eurograph. Symp. Comput. Animation*, San Diego, CA, USA, 2003, pp. 239–244.
- [15] K. Pullen and C. Bregler, “Motion Capture Assisted Animation: Texturing and Synthesis,” *ACM Trans. Graph.*, vol. 21, no. 3, July 2002, pp. 501–508.
- [16] N. Al-Ghreimil and J.K. Hahn, “Combined Partial Motion Clips,” *Int. Conf. Central Europe Comput. Graphic, Vis. Comput. Vis.*, Plzen, Czech Republic, 2003, pp. 9–16.
- [17] L. Ikemoto and D.A. Forsyth, “Enriching a Motion Collection by Transplanting Limbs,” *ACM SIGGRAPH/Eurographics Symp. Comput. Animation*, Grenoble, France, 2004, pp. 99–108.
- [18] G. Ashraf and K.C. Wong, “Generating Consistent Motion Transition via Decoupled Framespace Interpolation,” *Comput. Graph. Forum*, vol. 19, no. 3, Sept. 2001, pp. 447–456.
- [19] R. Heck, L. Kovar, and M. Gleicher, “Splicing Upper-body Actions with Locomotion,” *Comput. Graph. Forum*, vol. 25, no. 3, Sept. 2006, pp. 459–466.
- [20] M. Oshita, “Smart Motion Synthesis,” *Comput. Graph. Forum*, vol. 27, no. 7, Oct. 2008, pp. 1909–1918.
- [21] P. Glardon, R. Boulic, and D. Thalmann, “Robust on-Line Adaptive Footplant Detection and Enforcement for Locomotion,” *Vis. Comput.*, vol. 22, no. 3, Mar. 2006, pp. 194–209.
- [22] W.H. Press et al., “*Numerical Recipes in C++: The Art of Scientific Computing*,” New York, NJ, USA: Cambridge University Press, 2002.
- [23] Y. Kim and M. Neff, “Automating Expressive Locomotion Generation,” *Trans. Edutainment VII, LNCS*, vol. 7145, 2012, pp. 48–61.
- [24] C. Rose et al., “Efficient Generation of Motion Transitions Using Spacetime Constraints,” *ACM SIGGRAPH*, New Orleans, LA, USA, 1996, pp. 147–154.
- [25] K. Shoemake, “Animating Rotation with Quaternion Curves,” *ACM SIGGRAPH*, San Francisco, CA, USA, 1985, pp. 245–254.
- [26] J.-H. Yoo and M.S. Nixon, “Automated Markerless Analysis of Human Gait Motion for Recognition and Classification,” *ETRI J.*, vol. 33, no. 2, Apr. 2011, pp. 259–266.



Yejin Kim received his BS degree in computer engineering from the University of Michigan, Ann Arbor, USA, in 2000. He received his MS and PhD degrees in computer science from the Korea Advanced Institute of Science and Technology, Daejeon, Rep. of Korea, in 2003 and University of California, Davis, USA, in 2013, respectively. Currently, he is working for the SW·Content Research Laboratory, Electronics and Telecommunications Research Institute, Daejeon, Rep. of Korea, as a senior research scientist. His research interests include 3D character animation and authoring tools in computer graphics.



Myunggyu Kim received his BS and PhD degrees in physics from Seoul National University, Rep. of Korea, in 1989 and the University of Maryland, College Park, USA, in 1994, respectively. He is currently working for the SW·Content Research Laboratory, Electronics and Telecommunications Research Institute, Daejeon, Rep. of Korea, as a principal research scientist with interests in physical simulations and digital content.



Michael Neff received his MS and PhD degrees in computer science from the University of Toronto, Canada, in 1998 and 2005, respectively. He is an associate professor of Computer Science and Cinema and Technocultural Studies at the University of California, Davis, USA. His research focuses on character animation, gestures, and interactive techniques for computer graphics.