Context-Aware Motion Diversification
for Crowd Simulation*

Qin Gu and Zhigang Deng

Computer Science Department
University of Houston
Houston, TX, 77204, USA
http://www.cs.uh.edu

Technical Report Number
UH-CS-10-01

* This work has been accepted to *IEEE Computer Graphics and Applications* for publication on Feb 28, 2010.

Keywords: List Crowd Simulation, Motion Diversification, Motion Variety, Variety Realism, Character Animation, User Study.

Abstract

Traditional crowd simulation models typically focus on the navigational path-finding and local collision avoidance aspects. Relatively few existing efforts explore how to optimally control individual agents' detailed motions throughout a crowd. In this paper we propose a novel scheme for dynamically controlling motion styles of agents to increase the motion variety of a crowd. The central idea of our scheme is to maximize the style variety of local neighbors and global style utilization while maintaining the style consistency for each agent as natural as possible. Our scheme can serve as a complementary layer for most high-level crowd models to increase the variety realism. We show the flexibility and superiority of our scheme over traditional random motion style distribution through several experiment scenarios and user evaluations. To assist the runtime diversity control, an off-line preprocessing algorithm is also proposed to extract and stylize primitive motions from a motion capture database.

Context-Aware Motion Diversification for Crowd Simulation

Qin Gu and Zhigang Deng
Department of Computer Science, University of Houston, Houston, TX 77204
E-mails: {ericgu|zdeng}@cs.uh.edu

Abstract

Traditional crowd simulation models typically focus on the navigational path-finding and local collision avoidance aspects. Relatively few existing efforts explore how to optimally control individual agents' detailed motions throughout a crowd. In this paper we propose a novel scheme for dynamically controlling motion styles of agents to increase the motion variety of a crowd. The central idea of our scheme is to maximize the style variety of local neighbors and global style utilization while maintaining the style consistency for each agent as natural as possible. Our scheme can serve as a complementary layer for most high-level crowd models to increase the variety realism. We show the flexibility and superiority of our scheme over traditional random motion style distribution through several experiment scenarios and user evaluations. To assist the runtime diversity control, an off-line preprocessing algorithm is also proposed to extract and stylize primitive motions from a motion capture database.

Index Terms

Crowd Simulation, Motion Diversification, Motion Variety, Variety Realism, Character Animation, User Study.

I. Introduction

Large-scale crowd simulation has been widely used in computer games, movies, virtual training, and education applications. As one of the popularized schemes, agent-based crowd simulation considers the properties of each individual (agent) separately at every time step, hence it can produce highly realistic simulation of path navigation, cognitive reaction, collision avoidance, cognitive reaction, and animation control.

An agent-based crowd simulation system can be conceptually regarded as a three-layer hierarchy. The highest layer provides the navigation waypoints by path-finding and decision-making. The intermediate layer computes high-level motion information for every update (i.e., time step) through perception rules or social forces to achieve collision avoidance and collision response. The lowest layer can be described as a motion generator driven by parameters from the higher levels that visualizes the detailed animation of each agent. Within a crowd, these three layers collectively inform each agent the answers of the following three questions: (a) where is the final target? (b) where is the next step? and (c) how to perform the motion in the next step? While numerous approaches have been proposed for global navigation and local perception, relatively few research efforts have been focused on controlling agents' detailed motions throughout a crowd, for example, how to efficiently control the motion diversity of a simulated crowd.

In this paper, we propose a novel scheme to control the diversity of motion styles in agent-based crowd simulations, given a fixed number of motion styles. In this writing we use *"motion type"* to represent the general category of character motions such as walking, running and waiting. The term *"motion style"* refers to the variations within each motion type. In most crowd simulation models, a large number of agents will perform the same motion type such as walking. The ideal way is to assign a unique motion style to each agent during a certain time period, since individuals in a real-world crowd have diverse motion styles based on their distinctive personalities. However, this extremely high diversity appears prohibitive in terms of both computation and resource costs. Therefore, the problem of improving the motion diversity (or variety) of a crowd is transformed to *how to make the crowd look diversely plausible given a limited number of available motion styles*.

Previous crowd simulations usually specify a single motion style for a certain motion type. To increase the motion diversity, a random distribution of several motion styles is adopted in some applications. However, even the latter solution does not guarantee the motion diversity of the crowd since most likely it will lead to the local clustering or over-distribution of certain motion styles. Consequently, the cloned styles can be easily detected by audience. Our proposed motion diversity control scheme is designed to maximize both local and global diversity of motion style distribution, and special care is taken to ensure the realism and smoothness of motion transitions between different motion styles. Specifically, we use an offline data-driven method to extract and stylize primitive motions based on their kinetic energy characteristics. At runtime, the motion style of each agent at every time step is dynamically computed based on the information of its local neighbors, current style, and the global style distribution. The generality, robustness, and effectiveness of our approach was demonstrated by applying it to both cyclic motions such as walking and running, and acyclic motions such as fighting and waiting motions.

The main contribution of this work includes: (1) An efficient scheme to dynamically control the diversity of motion styles for agent-based crowd simulations, and (2) a data-driven motion stylization scheme from a motion capture dataset.

II. Related Work

In recent years, increasing attention has been attracted to agent-based models that simulate sophisticated global path-planning and local collision dynamics of each crowd member separately. Among them, force-based models and its various extensions such as [1] apply repulsion and tangential social forces to drive interactions between individual agents or sub-groups. Following the Reynolds's seminal work [2] for generating steering behaviors for flocks, herds, and schools, rule-based crowd simulation has been extensively studied to achieve highly realistic human behaviors in complicated environments. In addition, the widely known "motion graphs" algorithm [3] has also been introduced to retrieve and playback the appropriate motion data in crowd simulations.

Visual variety is an important factor to affect the overall perception of many crowds simulation scenarios, such as a street with high density pedestrians. Due to the computation and resource limitations, most real-time simulation systems have to employ repeated agent appearance or motion patterns for efficient runtime performance while sacrificing the crowd diversity to a certain extent. A number of approaches have been proposed to enrich the appearance variety among agents such as re-coloring textures for different body parts at runtime [4], modulating illumination maps [5], manipulating combinations of personal accessories and scaling body skeletons for different body heights [4]. In addition, Johansson [6] had investigated the problem of visual perception of biological motions and found that a 10-12 moving dot representation was adequate to evoke a compelling impression of human motions (e.g., locomotive human motions).

The most related work to our proposed scheme is the perception study of crowd variety by McDonnell and her colleagues [7]. In their study, several design implications and rules of crowd variety were derived such as how appearance clone, motion clone, and their combinations can affect the perceived variety of a crowd. McDonnell and her colleagues further evaluated the perceptual influences of different parts of a human body in a crowd [8]. These evaluations prove the effectiveness of adding appearance variety and illustrate that adding motion styles can also contribute to disguising clone effects. Compared with diversifying an agent appearance which has to be generated at the beginning and fixed during the simulation, dynamically changing motion styles of individual agents over time, to a certain extent, is less likely to be visually detected, hence having more controllability at run-time.

Researchers also investigated how to generate motion variations based on a given character motion dataset. For example, Lau *et al.* [9] developed a new dynamic Bayesian Network model to evaluate motion variations with a fast speed in both temporal and spatial domains. However, it is still impractical to directly apply these single-character animation techniques to a large-scale crowd with thousands of agents, since it is too computationally costly for real-time applications. To this end, we focus on a general approach to efficiently and adaptively enhance the motion variety and thus visual realism of a crowd, given a limited set of available motion styles.

III. Pipeline Overview

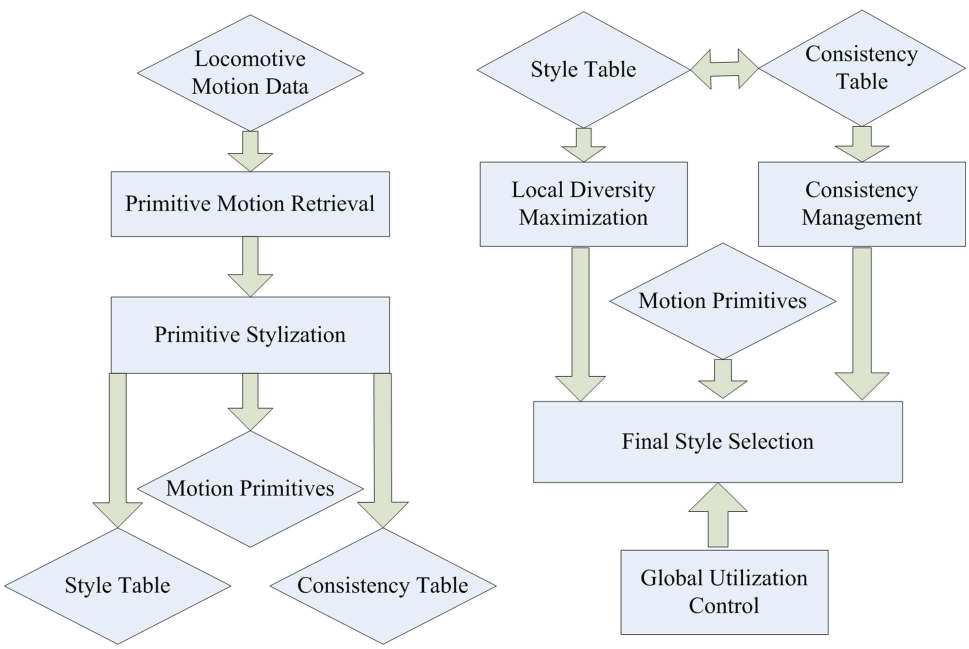

Fig. 1. The pipeline overview of our approach. The left shows the pipeline of our off-line motion stylization process, and the right shows the pipeline of the runtime motion diversity control scheme.

Figure 1 shows the pipelines of two main components of our approach. In the offline preprocessing step, we first segment and extract different primitive human motions from a motion capture database. These motions include cyclic walking and running motions, and acyclic waiting and fighting motions. Then, a novel stylization process parameterizes and sorts the obtained primitive motions based on kinetic energy. Style variation tables and compact consistency tables are generated for the runtime query. At runtime, the animation layer first retrieves feature vectors from higher perception layers of a crowd simulation system to decide the *motion type* and velocity of each agent. Then, our novel motion diversity control scheme selects proper *motion styles* for individual agents at each update (time step).

Our motion diversity control scheme is three-fold, based on the following simulation premises: (a) The chosen motion style of an agent should maximize the variety of local style distribution of its neighbors, so that the same or highly similar motion styles are not clustered together; (b) the chosen motion style of an agent should maximize the diversity of global motion style utilization; and (c) for certain motions such as cyclic walking, the chosen motion style of an agent should be consistent as much as possible with his/her current style to prevent unrealistic sharp changes of motion.

IV. Offline Motion Preprocessing

To assist the runtime motion diversity control, primitive motions with associated style information are generated at the offline preprocessing step. Although our motion diversity control scheme is independent of motion type, without the loss of generality, this work only considers walking, running, fighting and waiting motions since they are commonly used character motions in crowd simulations such as battlefields and urban streets.

We choose motion capture data due to its accuracy and realism. However, capturing a large motion capture database for every crowd simulation project is not always practical in reality; hence, we propose a data-driven method to extract and stylize primitive motions from a publicly available motion capture database, e.g., the well-known CMU motion capture database (http://mocap.cs.cmu.edu).

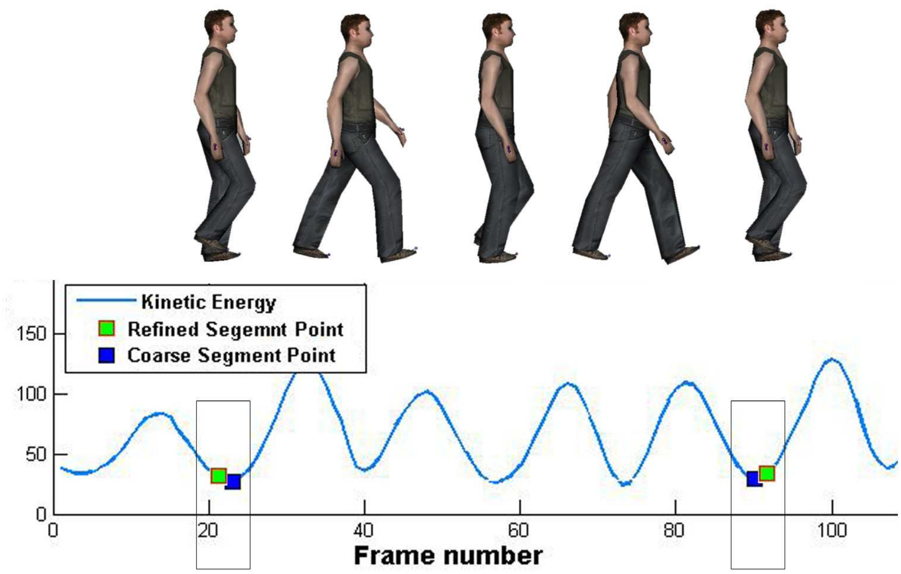

Fig. 2. Primitive motion extraction. (Top) A full cycle walking. (Bottom) Kinetic energy segmentation (its refining window size is 7).

A. Primitive Motion Extraction

Retrieving characteristic motions from a large motion database has been extensively studied during the past several years. Both semantically and numerically based retrieval approaches achieve an impressive accuracy of classifying different types of motions. Due to the high performance requirement of large-scale crowd simulations, we need an efficient and concise feature vector to characterize motion styles.

A number of previous efforts have been attempted to analyze, decompose, and quantify human motions [10], [11]. For example, Troje [10] presented an efficient framework to decompose walking motions to a low dimensional representation for analysis and synthesis purpose. However, whether this framework can be soundly and robustly applied to other types of human motions (not limited walking motions) was not established. Ren et al. [11] explored statistical models to quantify the naturalness of various human motions. But their approach was focused on the qualitative judgment of natural versus unnatural aspects of human motions. Therefore, their approach is not able to produce a quantitative feature vector for characterizing each motion style. In this work, inspired by the distance function proposed by Onuma et al. [12], we use the instant kinetic energy computed from joint angular velocity and joint moment of inertia to retrieve primitive motions and adopt the mean kinetic energy to stylize each primitive. Compared with original joint angle motion data, instant kinetic energy is invariant to reciprocal limb motions. For instance, in a walking motion, the raw movement angle of one leg will be neutralized by the negative movement of the other leg.

Proper segmentation is critical for cyclic walking and running motions. We retrieve primitive motion segments that start from the single foot contact state and consist of a full cycle of motions (Fig. 2) based on the following observations: (a) The starting pose and ending pose of a full motion cycle are most similar so as to optimize the runtime blending, (b) the switching-nature pose among real-world walking, running and standing is always the single foot contact pose, and (c) the single foot contact pose is more common than any on-the-fly poses regardless motion styles.

Primitive motion segments can be roughly identified by analyzing the kinetic energy trajectory of the entire human body since a locomotive motion exhibits highly cyclic patterns. Fig. 2 shows the low-pass filtered kinetic energy curve of a walking motion. We use the method proposed by Onuma et al. [12] to compute the moment of inertia of each joint. Unlike the COM (Center Of Mass) trajectory, the kinetic energy values fall into a local minimum on both constrained poses (e.g., the single foot stage) and unconstrained poses (e.g., the double foot stage for walking and the flying stage for running). This helps us to unify a segmentation solution for both walking and running motions. A full locomotion cycle can be segmented by starting from a single foot contact stage (a local minimum and foot contact with ground) and keep searching for five consecutive local valleys. If a bursting spike (sharp turn motion) or a number of continuous near-zero value (static pose) is encountered before the 5^{th} local valley, the current search is reset.

The above procedure provides us a coarse segmentation for cyclic walking and running motions. To further ensure the seamless transition between extracted motion segments, the obtained coarse segmentation is refined by applying a small window-based check around the first and last frame (Fig. 2). We compute the optimal (i.e. the smallest distance) frame pair as the final segmentation points using the metric proposed by Kovar et al. [3]. In this work, the size of the check window is experimentally set to 7 and 5 for walking and running motions, respectively. The refining process results in two types of motion cycles: one starting from the left foot and the other starting from the right foot. In this work, we only use the right foot cycle, since the left foot cycle can be easily obtained by swapping the halves of the right foot cycle. On the other hand, most acyclic motions, such as fighting and waiting motions, do not exhibit repeated motion patterns. Thus, acyclic primitives are retrieved by detecting long foot contact period with a threshold of kinetic energy change. We found this simple solution is sufficient to extract waiting and fighting motions with comparable style variations.

B. *Motion Stylization*

In order to control the motion diversity of individual agents, we need to design a metric to quantitatively measure the difference between a pair of primitive motion styles. Our stylization process first categorizes unlabeled primitive motions into different types, and then further classifies the motions with the same type into different styles.

The logarithm of the mean kinetic energy [12] is an effective metric for our purpose since it is independent of the length of the motion. We empirically choose the following 2D feature vector (Eq. 1), composed of the logarithm of the mean kinetic energy of upper/lower body parts, for motion clustering and classification.

$$(log(E_{upperbody} + 1), log(E_{lowerbody} + 1)) \tag{1}$$

Within each motion type, this metric is also used to quantify the style variations. In our experiment, this metric generally resulted in a perceptually sound motion style ranking.

To accelerate runtime motion selection, for each motion type, we also generate a style variation table (Fig. 4) with a size $s \times s$, where s is the number of styles for certain motion type and the value of each cell is the Euclidean distance between two motion styles. This data structure will be used as a look-up table in maximization of local motion diversity (described in Section V-A). Furthermore, a consistency table pre-registers a number of highly similar primitive motions for a specific style as possible candidates in style selection (discussed in details in Section V-C). To prevent foot-sliding, at runtime we need to align the original speed of primitive motions with the speed computed from the high level layers of a crowd simulation system. In this work, the original speed of a primitive motion is calculated by averaging the horizontal speed of its root, and its runtime resampling factor is computed as the ratio between the two speeds.

V. Dynamic Motion Diversity Control

The diversity or variety of a crowd is often categorized to appearance variety and motion variety. Having the same appearance or motion on a large number of individual agents in a crowd will make the simulation unrealistic. Appearance variety can be achieved through generating different 2D textures for the same 3D model. However, synthesizing realistic new motion styles for each agent is not practical for a large crowd in real-time applications. Previous studies also showed that adding appearance variety would not help to increase the diversity of motion styles in a crowd [7].

Our motion diversity control scheme explores an intelligent way to dynamically distribute the given limited motion styles into a large crowd. Given a specific motion type from high-level crowd simulation modules, for an agent p at time t, the

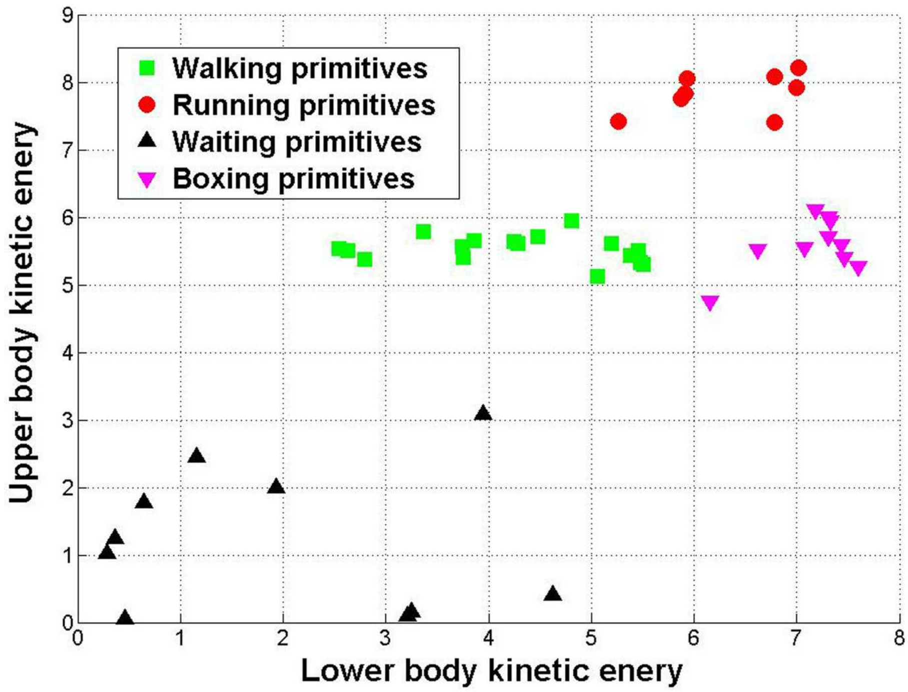

Fig. 3. Feature vector distribution in 2D lower-body and upper-body kinetic energy space.

	1	2	3	4	5	6	7
1	0.00	1.73	3.46	5.20	6.93	8.66	10.39
2	1.73	0.00	1.73	3.46	5.20	6.93	8.66
3	3.46	1.73	0.00	1.73	3.46	5.20	6.93
4	5.20	3.46	1.73	0.00	1.73	3.46	5.20
5	6.93	5.20	3.46	1.73	0.00	1.73	3.46
6	8.66	6.93	5.20	3.46	1.73	0.00	1.73
7	10.39	8.66	6.93	5.20	3.46	1.73	0.00

	1	2	3	4	5
1	1	2	3	4	5
2	2	1	3	4	5
3	3	2	4	1	5
4	4	3	5	2	6
5	5	4	6	3	7
6	6	5	7	4	8
7	7	6	8	5	9

Fig. 4. (Left) An example of walking style variation tables, and (right) an example of consistency tables (the row headers are motion style index number, and the column headers are the consistency ranking).

optimal motion style for next time interval, $S_p[t+1]$, is computed as the weighted combination of a local diversity function $D_p(S)$, a global utilization function $U(S)$, and a consistency management function $C_p(S)$:

$$S_p[t+1] = \arg\max_{S \subset R}(D_p(S)w_d + U(S)w_u + C_p(S)w_c) \quad (2)$$

Here S is a motion style candidate from the space R of all available styles of the expected motion type, w_d, w_u, and w_c are user-defined parameters to weight different control components. The style-updating interval is the length of each primitive motion described in Section IV-A.

A. Maximization of Local Motion Diversity

The maximization of local motion diversity is inspired by the optimal graph-coloring problem. However, computing an optimal k-coloring for a set of nodes is NP-hard. In addition, in a high density crowd, the number of neighboring nodes might be larger than the number of available motion styles. We therefore refine the selection criterion, that is, it should find the most different motion style compared with neighboring agents within the local field of interest. Fig. 5 shows a side-by-side illustration of this selection criterion. In this figure, gray-scale levels in the left panel indicate the kinetic energy representations of motion styles of agents, and the right panel shows the corresponding agent animations.

The local diversity function $D_p(S)$ for an agent p is computed as follows:

$$D_p(S) = \frac{1}{m}\sum_{q=1}^{m}\frac{StyleTable(S_q[t], S)}{NormalizedDistance_{pq}} \quad (3)$$

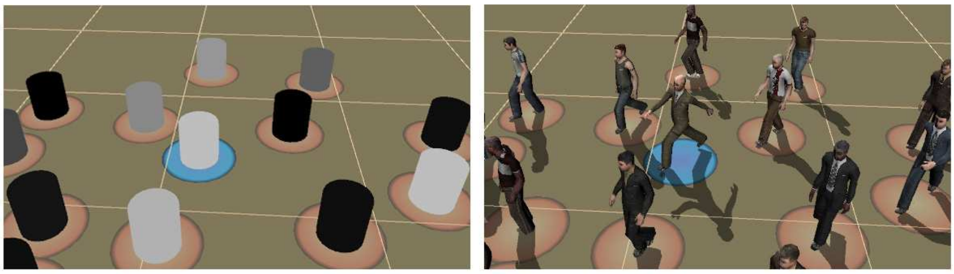

Fig. 5. The selected agent (blue) attempts to perform a walking style that is most distinguished from its neighbors (red). (Left) Kinetic energy representation of motion styles. (Right) The corresponding 3D agent animations.

Here m is the number of neighbors around the agent p, the numerator denotes the style difference between the candidate style S and current style of a neighbor agent q at time t, and it can be retrieved by looking up the style variation table constructed in Section IV-B. The $NormalizedDistance_{pq}$ between two agents p and q gives a closer neighbor with a higher weight (importance) and a more distant neighbor with a lower weight on style selection. Note the Euclidean distance between two agents in the world coordinate space cannot be directly used since the value $StyleTable(S_q[t], S)$ is computed in a stylization metric space. To avoid too large or too small value of $D_p(S)$, the distance in the world coordinate space is hence normalized by the average distance between the agent p and its m neighbors as follows:

$$NormalizedDistance_{pq} = \frac{distance(pq)}{\frac{1}{m}\sum_{i=1}^{m} distance(pq_i)} \tag{4}$$

Finding m nearest neighbors is a bottleneck for both high-level perception simulation layers and our local diversity maximization part due to its $\Theta(n^2)$ complexity, where n is the total number of agents in simulation environments. In this work, we address this issue by registering each agent into a discretized 2D grid at the beginning of every update and then only looking for nearest neighbors within the current agent's grid and 8 adjacent grids (Fig. 5). Note in the high-level perception level simulation, the neighbors are typically restricted to the area in front of the agent to mimic the vision angle restriction of humans (i.e. agents). In our local diversity control part, however, we consider the neighbors from all directions since our purpose is to disguise style clones from audience instead of agents. Moreover, only neighbors with *the same motion type* are taken into account since the style variation tables of different motion types are generated separately in Section IV-B.

B. Global Utilization Control

The selected motion style is expected to contribute to the optimization of global utilization distribution of all available styles. For a motion style S, its corresponding global utilization function $U(S)$ is computed as follows:

$$U(S) = targetNum(S) - currentNum(S) \tag{5}$$

Here $targetNum(S)$ is the expected occurrence number of the style S, and $currentNum(S)$ is the current occurrence number of the style S. $U(S)$ can be either positive or negative. A larger value of $U(S)$ means the crowd has none or only few replications of style S, while a smaller value of $U(S)$ indicates that the style S has already been "over-cloned" and thus will be repulsed by our diversity control model. $currentNum(S)$ can be simply obtained by keeping a style counter over time. $targetNum(S)$ is an empirical parameter derived from the distribution of global style utilization as follows.

$$targetNum(S) = P_S \times \frac{agentNum(T)}{styleNum(T)} \tag{6}$$

Here $\frac{agentNum(T)}{styleNum(T)}$ is the ratio between the number of agents with a motion type T and the number of available styles in the motion type T (assuming S is one of motion styles of the motion type T), and it represents the average number of style distribution on agents; P_S denotes the priority of the style S. By default, P_S is 1 for any style S. The strategy of our global utilization control is to maximize the utilization of every available motion style so that an approximately uniform distribution can be achieved. Fig. 6 compares the results of a 200-agents crowd with and without global utilization maximization. A crowd simulation with a random distribution of motion styles often leads to an unbalanced style distribution whereas our proposed scheme produces a near-uniform style utilization. In some special scenarios, certain styles are preferred than others. The preference can be achieved by increasing their priority values in Eq. 6.

C. Consistency Management

We also assume that an agent in a crowd likes to keep its motion style as much as possible. This is intuitive since if agents frequently switch very different motion styles, the entire crowd will appear unrealistic and unsmooth. Through the constructed

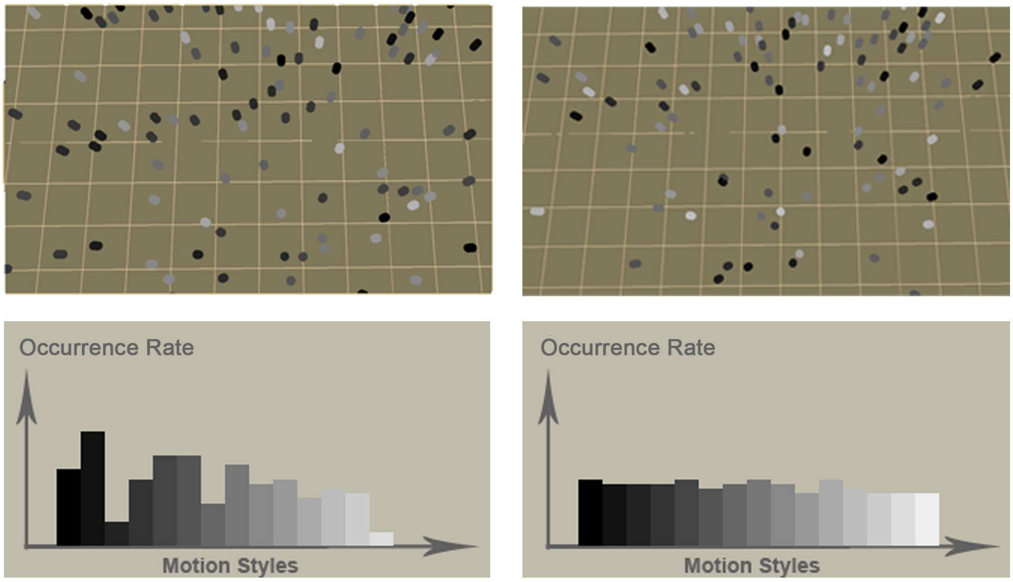

Fig. 6. (Left panel) Random distribution of motion styles, and (Right panel) our global utilization maximization.

style variation tables (Section IV-B), the consistency management function $C_p(S)$ for an agent p is computed as follows:

$$C_p(S) = \alpha_p \times (maxDistance - StyleTable(S_p[t], S)) \tag{7}$$

Here $StyleTable(S_p[t], S)$ represents the difference between current style of the agent p and a candidate style S, $maxDistance$ is the maximum distance/value in the style variation table of the specific motion type that S belongs to, and α represents whether the expected motion type is the same as current motion type (Eq. 8). We only consider the consistency between two styles in the same motion type since the change of motion type is much more obvious and normally controlled by high-level perception simulation layers.

In order to maintain the motion smoothness while maximizing the crowd variety, we adopt the widely-used Level of Detail (LOD) concept into our crowd diversity control scheme, that is, agents that are closer to the viewing camera are assumed to attract more attention from audience. Specifically, the consistency management function $C_p(S)$ is only applied to the agents in the visible range of the viewing camera; and if an agent is in the visible range, the influence/weight of the consistency management function is inversely proportional to the distance between the agent and the viewing camera (Eq. 8).

$$\alpha_p = \begin{cases} 0 & if \quad T_p[t+1] \neq T_p[t] \\ & or \quad Agent\ p\ is\ out\ of\ view \\ \dfrac{threshold}{dist(p)} & otherwise \end{cases} \tag{8}$$

Here $threshold$ is a scaling parameter.

Note that $D_p(S)$, $U(S)$, and $C_p(S)$ in Eq. 2 are functions of a style parameter S, which means in order to find the optimal style, we need to traverse the entire stylization space R of the same motion type. As such, we build an additional consistency table to reduce the size of the search space R based on the following key observations (Fig. 4):

- When a rich number of styles are used, traversing the entire style space at every update step is not efficient for a large-scale crowd simulation.
- Under the consistency constraint, candidate styles with a large difference from current style are rarely selected.

For each motion style in the style variation tables (Section IV-B), we first sort other styles within the same motion type by kinetic energy distance in an ascending order. Only the indices of the first r styles are stored in the consistency table. For any specific motion style, the stored candidates are therefore the r most similar styles (with a little higher or lower kinetic energy) starting from itself (zero variation). The right panel of Fig. 4 shows an example of the constructed consistency tables.

At runtime, an agent only traverses its closest r candidates (stored in the consistency table) at each update step to search for potential style transition targets. On the other hand, if users want to have more different styles on an agent, the agent should choose the r^{th} closest candidate style at current update step and then gradually move towards more distant style in next update step. For example, considering the consistency table with $r = 5$ in Fig. 4, in order to switch current #1 motion style to the target #7 style, our scheme will first transit #1 style to a middle #5 style and then reach #7 style at next update. We also require motion styles have a roughly even distribution in the stylization space (Fig. 3) to ensure that each style has at least one path to any other style in the consistency table. Thus, we remove redundant primitive motions (too close to or too distant from other primitive motions) before the style variation table is generated. Also, with the consistency table, $maxDistance$ in Eq. 7 can be empirically set to the r^{th} closest candidate distance.

VI. Results and Evaluation

To test and evaluate our diversity control scheme, we applied it to a number of crowd scenarios generated by high-level crowd simulation models [1], [2]. We also extracted 15 walking, 10 running, 10 fighting, and 10 waiting primitives from 11,138-frame motion sequences of the publicly available CMU motion capture database. Runtime performance and a perceptual user study are also reported in this section.

A. Experiment Scenarios

Fig. 7. Flocking Boids using our diversity control. Agents' motion styles converge to a stable stage after several update steps.

1) **Flocking** is a typical crowd behavior. Similar locomotion patterns and targets are observed among flock mates. We adopt the $Boids$ model [2] where each agent is driven by three steering behaviors: *separation*, *alignment* and *cohesion* (Fig. 7). Switching walking styles frequently will be easily detected in this model as the unnatural effect since relative speeds and orientations among agents are consistent. Motion style consistency in this case is given a higher weight/priority over local diversity and global utilization control. Its style distribution will converge to a stable stage after several update steps.
2) **Crowded Town** scenario is used to test various types of cyclic motions using the $HiDAC$ model proposed by Pelechano et al. [1] for both low density and high density crowds (Fig. 8). The walking crowd found the balance among the three control terms in Eq. 2. Meanwhile, switching running styles frequently in a panic situation does not produce obvious annoying effects. Based on these observations, we assume that motion style consistency has a lower influence on high frequency cyclic motions, and the local diversity maximization and global utilization control should be applied with higher weights.
3) **Frozen Land** scenario shows how our diversity control scheme can be applied to acyclic fighting, waiting and watching motions (Fig. 8 c, d). In contrast with cyclic locomotions, acyclic motions usually benefit from a higher weight on local diversity, a higher weight on global utilization, and a lower weight on consistency management to achieve non-repetitive motion patterns.
4) **Military March** shows the flexibility of our method through locally manipulating the global utilization control. While the default global utilization control tends to unify the style distribution, specific motion styles can be clustered by increasing their priority values P_S in Eq. 6 for certain selected agents. This is particularly useful in simulating formation crowds. Please refer to the accompanying demo video for its animation results.

The above four scenarios showcase the influence of weight tuning in our motion diversity control scheme. The weight parameters (Eq. 2) used in these experiment scenarios are shown in Table I.

B. Complexity and Performance

Performance is a critical issue for an agent-based crowd simulation system since the system has to update every agent individually at each time step. An unoptimized version of our motion diversity control has a computational complexity of

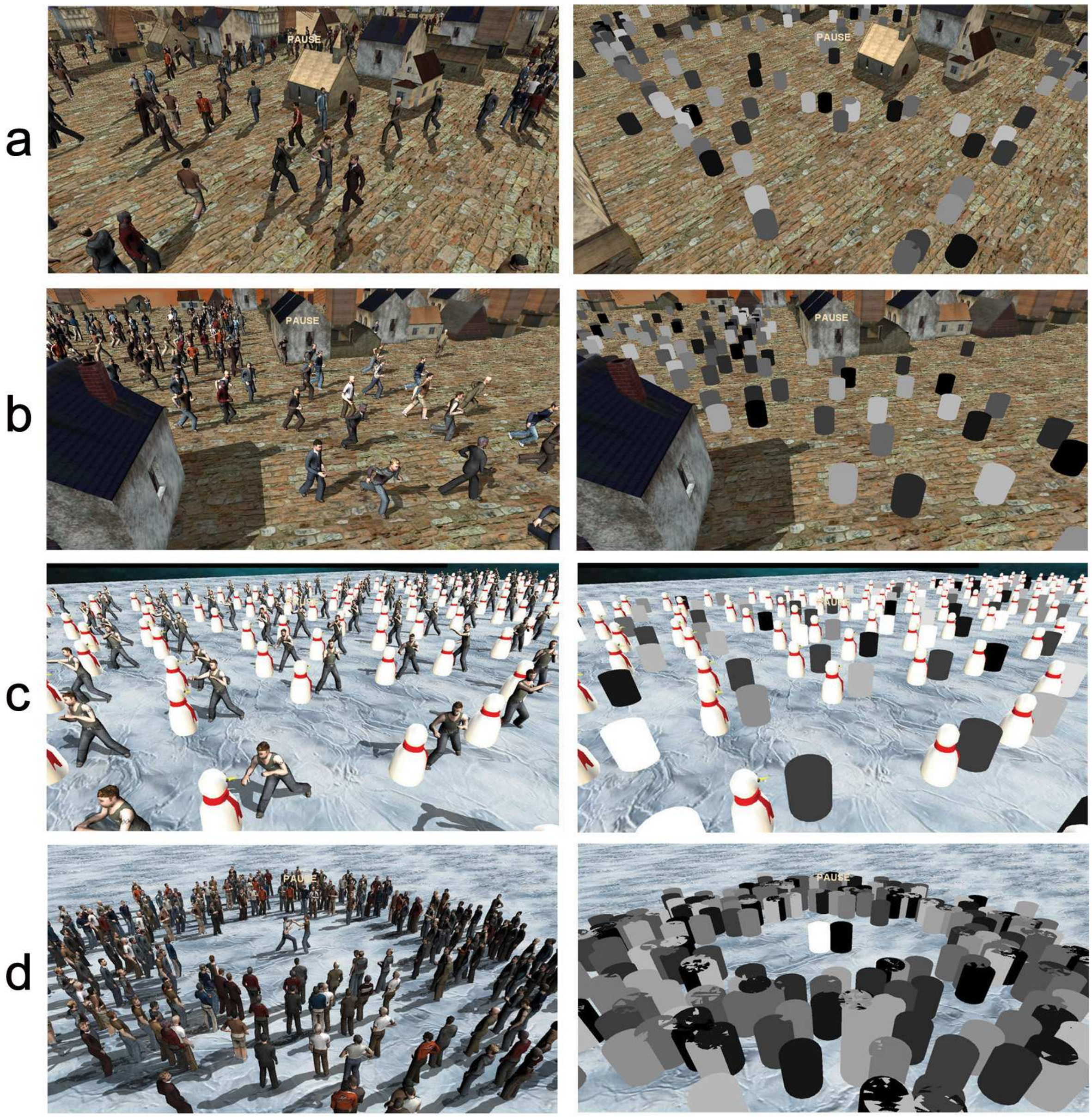

Fig. 8. Our motion diversity control is applied to different scenarios simulated by the HiDAC crowd model [1] (from top to bottom): (a) urban street, (b) panic evacuation, (c) fighting training, (d) street entertainment.

Simulation Scenario	Local Diversity (w_d)	Global Utilization (w_u)	Consistency (w_c)	Style Priority (P_S)
Flocking	1.0	1.0	5.0	1.0
Urban Street	1.0	1.0	1.0	1.0
Panic Evacuation	5.0	5.0	1.0	1.0
Fighting training	10.0	5.0	1.0	1.0
Street entertainment	10.0	5.0	3.0	1.0
Military March	1.0	10.0	5.0	10.0

TABLE I
WEIGHT PARAMETERS FOR LOCAL DIVERSITY MAXIMIZATION, GLOBAL UTILIZATION AND CONSISTENCY MANAGEMENT IN EQ. 2. IN THE MILITARY MARCH CASE, P_S IS 10.0 ONLY FOR MARCHING WALKING STYLES.

$n^2 \times s$, where n^2 is the cost of finding neighbors shared by higher-level perception layers and our diversity control, and s is the total number of motion styles. In our optimized implementation, n^2 is reduced to n by registering agents into a discretized 2D grid system at each frame. The s value is decreased to a small constant r through the introduced consistency table, as described in Section V-C.

An off-the-shelf PC with 2.4 GHz CPU, 2GB memory and NVidia GeForce 260 was used in our experiment. Using articulated 3D human models (800 to 1000 polygons) driven by high quality motion capture data with 30 joints (62 DOF), we are able to simulate up to 500 agents with 30 frames per second. To test the computation overhead, we compared the average Frames-Per-Second (FPS) when the random motion style distribution scheme (complexity is $\Theta(1)$) and our scheme are used, respectively (Table II). The high-level simulation used in the experiment is the HiDAC model [1] with discrete grid resolution of 50×50. Table II shows that our motion diversity control scheme only added a small overhead on top of the high level crowd simulation system.

crowd size	original fps (random distr.)	new fps (our scheme)	computation overhead
100 agents	123.5	122.9	0.7%
200 agents	97.7	97.2	0.5%
300 agents	60.1	59.8	0.5%
400 agents	39.4	39.0	1.0%

TABLE II
PERFORMANCE STATISTICS: THE OVERHEAD IS EVALUATED BY COMPARING THE UNLIMITED FRAME RATES BETWEEN THE ORIGINAL HIDAC MODEL AND THE HIDAC MODEL AUGMENTED WITH OUR MOTION DIVERSITY CONTROL SCHEME.

C. Perceptual Evaluation

Evaluating a simulated crowd numerically is a rather complicated problem since it highly depends on users' subjective perception. McDonnell et al. [7] provide an in-depth study on multiple factors that may affect the detection time of cloned motions from a crowd, including appearance, gait style, and the number of clones. In their work, the cloned motions are randomly distributed among the crowd without a control scheme. Therefore, our experiment goal is to answer the following usability question: compared with the random style distribution, does our diversity control scheme make motion clones harder to detect when the same number of motion styles is used?

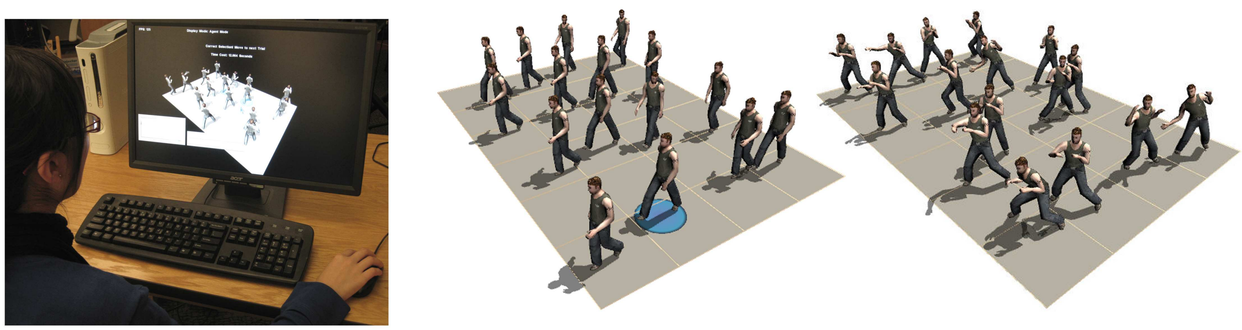

Fig. 9. Perceptual user study: (Left) Experimental interface, (Middle) cyclic walking motion, and (Right) acyclic fighting motion. The positions of 16 agents are uniformly generated plus a small random offset for each trial.

To minimize the influence from other simulation layers (e.g., navigation and perception levels), we fixed the positions and orientations of 16 testing agents with the same appearance. Cyclic walking and acyclic fighting primitive motions shown in section VI-A were used in the experiment. All agents are facing the same direction and not colliding with each others (Fig. 9). Different from the work of [7] that uses single clone style repeatedly, we attempted to simulate more specific crowd situations to allow multiple style clones multiple times, that is, given a limited number of motion styles, we let the random or our diversity control scheme to determine which and where a motion style should be applied. The upper-limits of available styles for each trial are set to 2, 4, 6, 8, and 10. These motion styles were chosen from the collected 15 walking and 10 fighting primitive motions at a random order, respectively.

In this study, the total trial number for each participant is 5 pools * 2 diversity schemes * 2 motion types = 20. 14 naive participants (12 Males, 2 Females) took part in the study. It should be noted that most of the participants had little crowd simulation background. Since it was difficult to find all clone pairs within a reasonable time frame, we asked each participant to pick one clone pair as quickly as possible. In order to eliminate the influences of different motion styles (some styles appear harder to identify than others) and fatigue issue as much as possible, the order of trials for all the participants was counter-balanced.

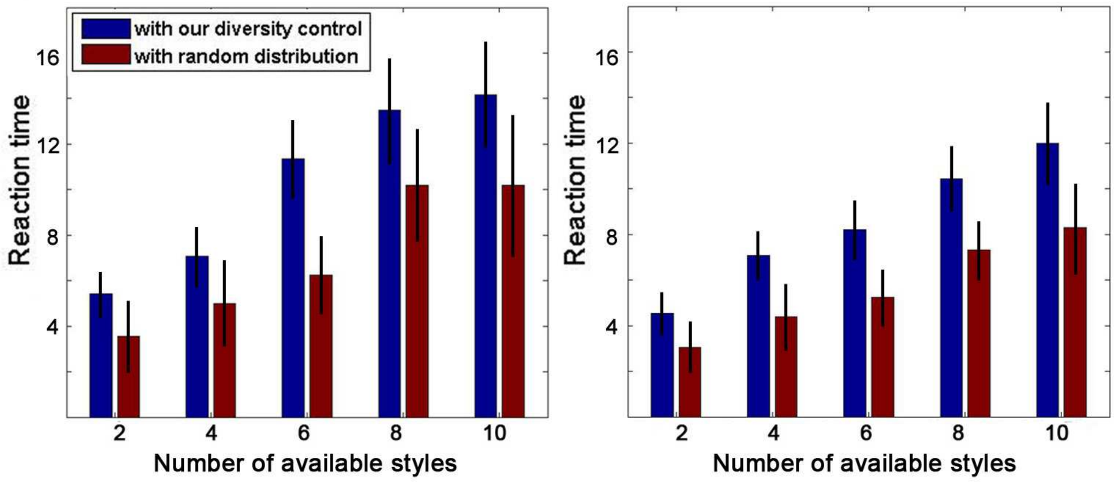

Fig. 10. Average response time and standard deviation of detecting the first pair of motion style clones: (left) cyclic walking motion, (right) acyclic fighting motion.

Two-way ANOVAs were used to analyze the reaction time of picking the first clone style in each trial. We found both the number of style upper-limit (cyclic motion: $F = 26.25$, p-value < 0.0039; acyclic motion: $F = 38.40$, p-value < 0.0019) and the diversity control scheme (cyclic motion: $F = 29.94$, p-value < 0.0054; acyclic motion: $F = 48.75$, p-value < 0.0016) are the main effects and there is no evident interaction between them. The first result (the number of the style upper-limit is one of the main effects) was consistent with the user study results reported by other researchers [7]. The other result (diversity control scheme is another main effect), combined with the average reaction time in both motion type conditions (Fig. 10), shows that our motion diversity control is more effective than the random distribution to disguise motion clones given the same number of available motion styles.

Although we did not evaluate the impact of agent orientation on motion clone detection, and allowed the participants to freely rotate the view during the study, we found most of the participants preferred to use a side view (Fig. 9), instead of a front view, to identify motion clones. This interesting observation is opposite to the case of detecting appearance clone reported in previous work [7]. One possible explanation is that most of the participants tried to identify different styles through the swing magnitude of limb motions that is more observable at the side view.

VII. Discussion and Conclusions

This paper proposes a novel scheme for dynamically controlling motion styles to increase the motion variety/diversity of agent-based crowds. The central idea of our scheme is to maximize the style variety of local neighbors and global style utilization while keeping the style consistency for each agent as natural as possible. The proposed scheme only requires high-level motion information (e.g., speed and motion type) computed from the navigation and perception layers of a crowd simulation system. As such, it can serve as a complementary layer for high-level crowd simulation models. Finally, we show the flexibility and superiority of our scheme over the traditional random motion style distribution through a number of experiment scenarios and a perceptual user study.

Although our motion diversity control scheme is independent of specific motion types, the off-line stylization process directly affects the final simulation results. Inappropriate stylization may cause jaggy effects in terms of consistency management, e.g., two styles have similar stylization value but not visually close. The proposed segmentation and stylization process generated sound results for the selected motions in this work. However, stylizing more complex human motions using compact feature vectors is still a challenging problem that remains to be further explored.

The proposed scheme has several limitations. In the current work, we do not use the transitions enclosed in the original motion capture data due to the following main considerations: (1) many inter-style transitions are not available in the original motion data, and (2) pre-generating all the possible motion transitions among all styles demands nontrivial extra overhead for a large crowd. For performance concern, we apply a spherical linear interpolation on agents' joint rotations and a linear interpolation on agents' translations to dynamically generate transitions at runtime. Since the consistency management described in Section V-C restricts the style to change in a "mild" way, we found the dynamically generated transition results are visually acceptable. However, if the motion continuity is strictly required for a certain agent, such as the main character in a computer game, a more sophisticated motion synthesis method would be necessary instead of replaying and interpolating existing motion styles.

Also, we use the average speed computed from original primitive motions as the reference for computing the runtime animation re-sampling rate. This may still result in minor foot-slidings for certain locomotive motions, since the speed of real-world humans is a periodical acceleration and deceleration procedure along the foot steps while the speed from high-level crowd simulation layers is typically a constant. In addition, the current algorithm considers every agent as the same type of person without variations in genders, personalities, heights, weights, or ages. Several recent efforts [4], [8] indicate the body shapes and even the motions of particular body parts will significantly influence the overall visual variety of a simulated crowd. We hence plan to explore more sophisticated motion style selection rules to take these factors into consideration in order to further enhance the visual realism of diversified crowds.

As a common limitation of data-driven methods, the simulation result of our scheme is limited to the capacity and variety of the given motion database. For example, if the number of available motion styles in our scheme is extremely low, motion clones will be easily detected. As a part of our future work, we plan to develop algorithms to synthesize new motion style variations on-the-fly based on the current optimal motion selection outcome, balancing visual realism and runtime performance.

In the current user study, we focus on how to effectively disguise motion style clones to increase the visual variety. An interesting future direction would be to investigate the user perception aspect resulted from the change of motion styles in a crowd. This visual perception may vary with particular motion styles, agent distance, and the number of crowd agents. The finding would give additional insights to crowd motion diversity control.

ACKNOWLEDGMENTS

This work was supported in part by NSF IIS-0914965, Texas NHARP 003652-0058-2007, and NSF CNS-0932272. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

REFERENCES

[1] N. Pelechano, J. M. Allbeck, and N. I. Badler, "Controlling individual agents in high-density crowd simulation," in *Proc. of SIGGRAPH/Eurographics Symposium on Computer Animation (SCA) '07*, 2007, pp. 99–108.

[2] C. W. Reynolds, "Flocks, herds, and schools: A distributed behavioral model," in *Proc. of SIGGRAPH '87*, 1987, pp. 25–34.

[3] L. Kovar, M. Gleicher, and F. Pighin, "Motion graphs," *ACM Transaction on Graphics*, vol. 21, no. 3, pp. 473–482, 2002.

[4] J. Maïm, B. Yersin, and D. Thalmann, "Unique character instances for crowds," *IEEE Computer Graphics and Applications*, vol. 29, pp. 82–90, 2009.

[5] S. Dobbyn, J. Hamill, K. O'Conor, and C. O'Sullivan, "Geopostors: a real-time geometry/impostor crowd rendering system," in *I3D'05: Proc. of ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, 2005, pp. 95–102.

[6] G. Johansson, "Visual perception of biological motion and a model for its analysis," *Perception and Psychophysics*, vol. 14, no. 2, pp. 201–211, 1973.

[7] R. McDonnell, M. Larkin, S. Dobbyn, S. Collins, and C. O'Sullivan, "Clone attack! perception of crowd variety," in *SIGGRAPH '08: ACM SIGGRAPH 2008 papers*. New York, NY, USA: ACM, 2008, pp. 1–8.

[8] R. McDonnell, M. Larkin, B. Hernández, I. Rudomin, and C. O'Sullivan, "Eye-catching crowds: saliency based selective variation," *ACM Transaction on Graphics*, vol. 28, no. 3, pp. 1–10, 2009.

[9] M. Lau, Z. Bar-Joseph, and J. Kuffner, "Modeling spatial and temporal variation in motion data," *ACM Transaction on Graphics*, vol. 28, no. 5, pp. 1–10, 2009.

[10] N. F. Troje, "Decomposing biological motion: A framework for analysis and synthesis of human gait patterns," *Journal of Vision*, vol. 2, pp. 371–387, 2002.

[11] L. Ren, A. Patrick, A. A. Efros, J. K. Hodgins, and J. M. Rehg, "A data-driven approach to quantifying natural human motion," *ACM Transaction on Graphics*, vol. 24, no. 3, pp. 1090–1097, 2005.

[12] K. Onuma, C. Faloutsos, and J. K. Hodgins, "FMDistance: A fast and effective distance function for motion capture data," in *Short Papers Proceedings of EUROGRAPHICS*, 2008.