

# Collective Crowd Formation Transform with Mutual Information based Runtime Feedback

Mingliang Xu<sup>†1</sup>, Yunpeng Wu<sup>1</sup>, Yangdong Ye<sup>‡1</sup>, Illes Farkas<sup>2</sup>, Hao Jiang<sup>3</sup>, and Zhigang Deng<sup>§4</sup>

<sup>1</sup>Zhengzhou University, P. R. China

<sup>2</sup>Hungarian Academy of Sciences, Hungary

<sup>3</sup>Institute of Computing Technology, Chinese Academy of Sciences, P. R. China

<sup>4</sup>University of Houston, USA

---

## Abstract

*This paper introduces a new crowd formation transform approach to achieve visually pleasing group formation transition and control. Its core idea is to transform crowd formation shapes with a least-effort pair assignment using the Kuhn-Munkres algorithm, discover clusters of agent sub-groups using affinity propagation and Delaunay triangulation algorithms, and apply subgroup-based SFM (social force model) to the agent subgroups to achieve alignment, cohesion and collision avoidance. Meanwhile, mutual information of the dynamic crowd is used to guide agents' movement at runtime. This approach combines both macroscopic (involving least-effort position assignment and clustering) and microscopic (involving SFM) controls of the crowd transformation to maximally maintain subgroups' local stability and dynamic collective behavior, while minimizing the overall effort (i.e., traveling distance) of the agents during the transformation. Through simulation experiments and comparisons, we demonstrate that this approach is efficient and effective to generate visually pleasing and smooth transformations and outperform several existing crowd simulation approaches including RVO, ORCA and OpenSteer.*

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: 3D Graphics and Realism—Animation; I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction Techniques

---

## 1. Introduction

In recent years aesthetic animation of group formation transformation has been increasingly used in feature animation films [BB10], video games [HSK12, FR12], mass performance rehearsal, tactical arrangements of players for sports teams training, and so on. Furthermore, group formation generation and control can also find its wide applications in many other scientific and engineering fields including but not limited to robot control, multi-agent systems, and behavioral biology.

A well-documented phenomenon is that both human motion and crowd dynamics are governed by the principle of least effort (PLE) [Zip49, FH85, GCC\*10]. For group formation control, the PLE principle implies that all individuals in

a group attempt to reach their pre-defined positions using the least total effort while maintaining their collective formation and movements, which shall typically lead to an aesthetic formation transform. In a mass performance, for example, individual performers are often asked to keep similar relative positions to their neighbors and minimize the overall formation change, in order to make a smooth and visually pleasing transition from current formation to the target formation.

To date limited efforts have been focused on automatically generating aesthetic group formation animations, to the best of our knowledge. [KLLT08, TYK\*09] can generate aesthetic transitions between key crowd formation configurations, with the involvement of non-trivial manual efforts. Gu and Deng [GD11b, GD13] proposed a sketching interface to specify various group formations. Henry et al. [HSK12, HSK14] proposed a single-pass algorithm to control crowds using a deformable mesh. Nevertheless, these approaches are primarily focused on either automatically sampling optimum agents from user-specified sketch input

---

<sup>†</sup> iexumingliang@zzu.edu.cn

<sup>‡</sup> ieydye@zzu.edu.cn

<sup>§</sup> zdeng4@uh.edu

or dynamically adapting to crowd environments. Generating visually pleasing transformations between different group formations is not the main focus of these approaches. Indeed, with existing approaches, it is non-trivial to achieve visually pleasing results (especially for large-scale groups) without considerable efforts.

Inspired by the above challenge, in this paper we propose a novel group formation transform approach to automatically generate a visually pleasing formation transformation, given source and target group formations. Specifically, we first convert the source and target formation shapes to Delaunay Triangulation (DT) representations, where each vertex represents the spatial position of an agent in its group. Based on the DT representations, we can quantify the effort of each agent during the transformation. By clustering agents into collective subgroups, we further extend shape manipulation techniques and the Social Force Model (SFM) to ensure collision-free movements for all the individual agents in real-time. We evaluate the effectiveness and robustness of our approach through various simulation experiments. In addition, by comparing it with three existing approaches (namely, RVO [VdBMLM08], ORCA [VdBGLM11] and OpenSteer [Ope04]) via several objective measures, we demonstrate that our approach can produce more visually pleasing and fluid crowd formation transformations than these approaches.

The main contributions of this work are two-fold: (1) We introduce a complete and robust solution that can automatically generate visually pleasing crowd formation transformations, including a novel formula to quantify the least-effort of an agent during a crowd formation transform and a new relative distance variance measure to divide agents to subgroups. Our experiments demonstrate the effectiveness, controllability and versatility of the proposed solution. (2) We introduce a number of quantitative and objective measures to evaluate the realism of crowd formation transformations, and the measures can be potentially used for various crowd simulation applications, not limited to our approach.

## 2. Related Work

Numerous crowd simulation and modeling approaches have been developed during the past several decades. Here we only briefly review recent efforts that are most related to this work. Interested readers can refer to the recent survey by Zhou et al. [ZCC\*10].

Rule-based crowd models are flexible to simulate various crowd agents through a set of delicately designed rules. The seminal work by Reynolds [Rey87] presented the concept of Boids that simulates flocks of birds and schools of fishes via several simple yet effective steering behavioral rules to keep the group cohesion, alignment and separation as well as avoid collisions between group members. Recently, Klotsman and Tal [KT11] provided a biologically motivated rule-

based artificial bird model, which produces plausible and realistic line formations of birds.

A distinct research line of crowd simulations is force-based model, originally developed from human social force study by Helbing and Molnr [HM95]. Later, it was further applied and generalized to other simulation scenarios such as densely populated crowds [PAB07], simulation of pedestrian evolution [LKF05] and escape panic [HFV00].

The PLE principle has been exploited for crowd simulation previously. Guy et al. [GCC\*10, GCLM12] introduced PLE-based algorithms to simulate large-scale crowds at interactive rates. Their approaches minimize the amount of effort each agent in a large crowd incurs by optimizing its bio-mechanically energy-efficient, collision-free trajectory. Their work shares certain similarities with our approach, but our approach need to handle more constraints simultaneously such as position constraint, maintenance of clusters and sub-groups, and the aesthetic aspect of group transitions. In addition, information theory [TKB11], continuum theory [TCP06], or synthetic vision [OPOD10] was also employed to simulate or evaluate large-scale crowds.

Recently, researchers also explored novel ways to animate large-scale crowds by cloning crowd motion from existing examples [GD11a, LCS\*12] or blending crowd animations [JCP\*10], to synthesize realistic individual trajectories with interactions at a microscopic scale [LJK\*12], and to generate heterogeneous crowd behaviors by adjusting the simulation parameters to emulate personality traits of individuals within a crowd [GKLM11]. In addition, interactive path planning and navigation algorithms for multi-agent systems [SAC\*08, GSA\*09, PvdBC\*11, KOOP11] (in particular, large-scale heterogeneous agents in complex dynamic environments) and effective long-range collision avoidance algorithms [GNL13] were also proposed.

Group formation control is a vital collective characteristic of many crowds. Existing approaches typically combine heuristic rules with explicit hard constraints to produce and control sophisticated group formations. For example, Kwon et al. [KLLT08] proposed a framework to generate aesthetic transitions between key crowd formation configurations. A spectral-based group formation control scheme [TYK\*09] was also proposed. However, in these approaches, exact agent groups' distributions at a number of key frames need to be specified by users. Gu and Deng [GD11b, GD13] proposed an interactive and scalable framework to generate arbitrary group formations with controllable transitions in a crowd. Henry and colleagues [HSK12, HSK14] proposed a single-pass algorithm to control crowds using a deformable mesh, and this approach can be used to control crowd-environment interaction and obstacle avoidance. In addition, they proposed an alternative metric for use in a pair assignment approach for formation control that incorporates environment information. However, these approaches either need non-trivial manual involvements [KLLT08, TYK\*09] or are

focused on intuitive user interfaces for formation control and interaction [GD11b, GD13, HSK12, HSK14]. In contrast, the focus of our approach is the automated generation and evaluation of aesthetic crowd formation transitions in an efficient manner. From this perspective, our approach is complementary to these previous approaches.

Compared to our previous approach [XWY12], we have extended this approach in the following ways: (1) Our previous work only uses the Kuhn-Munkras(KM) algorithm for balanced assignment of agent pairs in the source and target formations. In contrast, besides the similar balanced assignment of agent pairs, our new work also introduces a novel relative distance variance measure to divide agents to subgroups, and further extend the well-known social force model with subgroup constraints as a new movement control mechanism in order to maximally maintain the aesthetic realism of transformed crowds. (2) Our previous work uses mutual information as the only quantitative measure. In contrast, not limited as a static quality measure, our new work further utilizes the mutual information of the crowd at runtime and dynamically uses it to control crowd formation transform. (3) Our new work also introduces two additional objective measures (i.e., the stability of local structure and the extent of effort balancing) to quantitatively evaluate the crowd simulation results by our approach and other state-of-the-art approaches. Furthermore, not limited to evaluation of crowd formation transforms, we believe our introduced objective aesthetic measures can find their broad uses in various other crowd simulation applications.

### 3. Our Method

The pipeline of our approach is described below (also illustrated in Fig. 1). Specifically, it can be divided into four steps: data preparation, pair assignment, subgroup clustering, and movement control. The details of the four steps are described later in this section.

1. First, by converting the source and target formation shapes to Delaunay Triangulation (DT) representations [LS80], we quantify the effort of each agent during the transformation. We also formulate the least-effort as a linear assignment optimization problem (Section 3.2).
2. Then, we employ the affinity propagation algorithm [FD07] to cluster collective subgroups, where a new relative distance variances matrix is introduced for clustering (Section 3.4).
3. Finally, to reflect the collective movements in subgroups, each subgroup is manipulated as a whole using a DT graph. And, the Social Force Model (SFM) [HM95, HFV00, MPG\*10] is extended and optimized with subgroup constraints to ensure collision-free movements for all the individual agents in real-time (Section 3.5).

### 3.1. Data Preparation

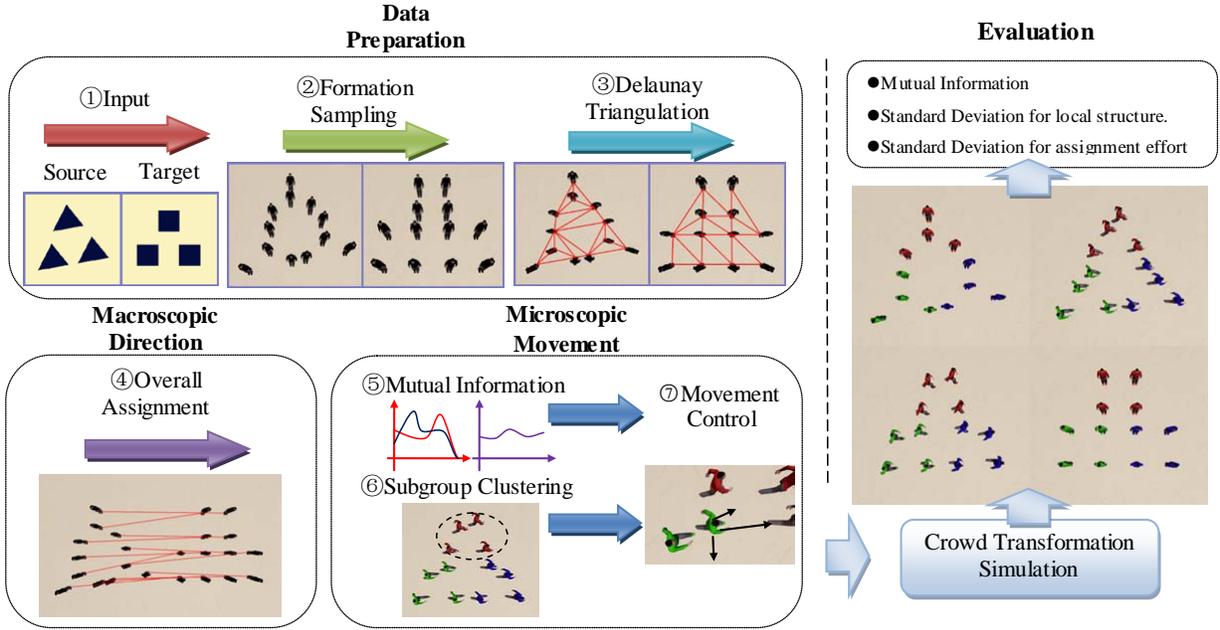
A crowd  $C$  is the set of agents, that is,  $C = \{a_i\}_{i=1}^N$ , where  $N = |C|$  is the number of agents in the crowd. In 2-dimension  $\mathbb{R}^2$  space, the position of an agent  $a_i$  is denoted as  $pos(a_i) = (x_i, y_i)$ . Let  $pos_{source}(a_i)$  and  $pos_{target}(a_i)$  be the source position and target position of  $a_i$ , then the sets of start points and target points are:  $S = \{s_i | s_i = pos_{source}(a_i), a_i \in C, i = 1 \cdots N\}$ , and  $T = \{t_i | t_i = pos_{target}(a_i), a_i \in C, i = 1 \cdots N\}$ .

The input data to our approach is the source and target formations, represented as user-specified bitmap files. The key step for data preparation is to sample a formation bitmap file with an appropriate density and a precise shape, so as to get the aesthetic visualization effect. In this work, we use a force-based formation sampling algorithm. It is inspired from the following physical principle: any particles existing in the environment have their natural properties such that the particles have a trend to distribute appropriately through the impact of repulse force with their neighbors; these interacting forces ensure the agents located neither too close nor too far.

The input files for our method are bitmaps represented as  $N_{len} \times N_{wid}$  pixel matrices, PIX, each of which is denoted as  $pix(x, y)$  ( $i = 1 \cdots N_{len}, j = 1 \cdots N_{wid}$ ).  $pix(x, y)$  is represented as RGB values  $(r_{i,y}, g_{i,y}, b_{i,y})$ . Then, a set of formation points that form the precise formation shape in the given bitmaps are filtered from PIX, which is defined as follows:

$$P_f = \{pix(x, y) | c_{min} \leq c_{i,j} \leq c_{max}, c \in \{r, g, b\}\} \quad (1)$$

In order to automatically use a user-specified number of agents to form a specified formation shape with a visually natural distribution, the sampling process from the formation shape is mainly divided into the following two stages: First, we use a simple way to tentatively sample the approximation of the parameterized number of formation points in the formation shape with a roughly even distribution. In this work, we choose appropriate sampling densities  $den_x$  and  $den_y$  to get a set of the sampled formation points denoted as the set of  $pix(x, y)$  where  $pix(x, y) \in P_f$ ,  $x \bmod den_x = 0$ , and  $y \bmod den_y = 0$ . However, the sampled result is very rigid, lacking the aesthetic effect. More importantly, it is difficult for this method to accurately sample a user-specified number of formation points. Therefore, we have to tune the number of sampled points to be exactly equivalent to the user-specified number by randomly deleting some sampled points or adding some formation points according to the Roulette Wheel Selection strategy. Then, a corresponding agent is located at the location of a formation point. It should be noted that other formation sampling strategies such as the formation coordinate based sampling method [GD11b, GD13] can be straightforwardly used in this step as well.



**Figure 1:** The pipeline of our crowd formation transform approach

### 3.2. Pair Assignment

After the sampling process, DT is employed to represent the relationship among adjacent agents. Let  $G(S, E_S)$  and  $G(T, E_T)$  be the DT of the point sets  $S$  and  $T$ , and  $E_S$  and  $E_T$  be the sets of edges in DT. In this work, we use  $G(S, E_S)$  and  $G(T, E_T)$  to denote the source and target formation graphs for the crowd  $C$ . Readers can refer to [LS80] for the details of how to construct DT from a set of points.

To generate the formation transformation, we need to pair the agents in the source formation with those in the target formation. This can be formulated as the problem of building a one-to-one correspondence between the source point set and the target point set. We can further formulate it as finding an optimal assignment in a weighted bipartite graph  $G_{bi}(V_{bi}, E_{bi})$ , where  $V_{bi} = S \cup T$  and  $S \cap T = \emptyset$ , and  $E_{bi}$  encodes the paired result.

The assignment from the source to the target is a one-to-one mapping function from  $G(S, E_S)$  to  $G(T, E_T)$ , denoted as

$$\Psi : S \rightarrow T. \quad (2)$$

For  $\forall s_i \in S, \forall t_k \in T$ ,  $s_i = pos_{source}(a) \wedge t_k = pos_{target}(a) \Leftrightarrow t_k = \Psi(s_i)$ .

In the matching process, we apply a novel match measure to effectively minimize the overall disorder including the variations of both time synchronization and local structure. For  $\forall s_i \in S$ , its nearest point set in  $G(S, E_S)$

is  $\{s_i\}_{np} = \{s_p | s_i s_p \in E_S, i \neq p\}$ , and in a similar way  $\{t_j\}_{np} = \{t_q | t_j t_q \in E_T, j \neq q\}$ , for  $\forall t_j \in T$ . Let  $\Delta E(i, j)$  denote the effort between  $s_i$  and  $t_j$ , which can be defined as follows:

$$\Delta E(i, j) = \lambda \|\vec{s_i t_j}\|^2 + (1 - \lambda) |\delta s_i - \delta t_j|^2, \quad (3)$$

where  $\|\vec{s_i t_j}\|$  is the length of the vector  $\vec{s_i t_j}$ ,  $\delta s_i$  and  $\delta t_j$  are the mean lengths of the edges associated with  $s_i$  and  $t_j$ , respectively, defined as follows:

$$\delta s_i = \frac{1}{|\{s_i\}_{np}|} \sum_{s_p \in \{s_i\}_{np}} \|\vec{s_i s_p}\| \quad (4)$$

$$\delta t_j = \frac{1}{|\{t_j\}_{np}|} \sum_{t_q \in \{t_j\}_{np}} \|\vec{t_j t_q}\| \quad (5)$$

In the above Eq. 3,  $\lambda$  balances the trade-off between the first term  $\|\vec{s_i t_j}\|^2$  that measures the disorder of the agents' speed variations so as to dynamically adjust the agents' time synchronization and the second term  $|\delta s_i - \delta t_j|^2$  that measures the disorder of local structure variations between the formation  $S$  and  $T$  to weaken the agents' collective behavior. Therefore, we minimize the overall disorder to obtain the optimal pair assignment as follows:

$$\min \left( \sum_{j=\Psi(i)} \Delta E(i, j) \right) \quad (6)$$

To this end, we apply the classical Kuhn-Munkres (KM) algorithm [Kuh55, Mun57] with the above measure to solve the pair assignment problem.

### 3.3. Mutual Information based Runtime Feedback

For an agent  $a \in C$ , we locally adjust its trajectory by applying social forces such as driving and repulse forces to navigate and avoid collisions. In the Social Force Model (SFM) [HM95, HFV00, MPG\*10], the resulting force  $F$  of the agent  $a$  is the sum of the following forces: the driving force  $f_{drive}(v_a)$ , the obstacles' force  $f_{obstacle}(pos(o))$ , and the repulse force with other agents  $f_{repulse}(pos(a), pos(b))$ , as described below.

$$F = f_{drive}(v_a) + f_{obstacle}(pos(o)) + \sum_{a \neq b} f_{repulse}(pos(a), pos(b)) + \xi_a, \quad (7)$$

$$f_{drive}(v_a) = \frac{1}{\tau_a} (v_a^0 e_a - v_a), \quad (8)$$

$$f_{obstacle}(pos(o)) = A e^{-\frac{|pos(a) - pos(o)|}{B}}, \quad (9)$$

$$f_{repulse}(pos(a), pos(b)) = \frac{A e^{-\frac{|pos(a) - pos(b)|}{B}}}{|pos(a) - pos(b)|}, \quad (10)$$

where  $v_a^0$  is the desired velocity of  $a$ ,  $e_a$  is the ideal heading,  $v_a$  is the current velocity,  $\tau_a$  is the reaction time,  $pos(a)$  and  $pos(b)$  are the position vectors,  $|pos(a) - pos(b)|$  is the distance between them, and  $|pos(a) - pos(o)|$  is the distance between agents and obstacles.  $A$  and  $B$  are two user-specified constants, and  $\xi_a$  is the fluctuation factor that adds random variations to the motion.

In this work we dynamically incorporate the mutual information based feedback to the SFM. As the preliminary, we briefly describe the concept of mutual information and how it can be adapted to measure crowd formations within the context of this work.

Mutual Information (MI) is a well-known concept in the field of information theory, and it is designed to quantify the mutual dependence between two random variables. In general, the MI of two discrete time-series variables,  $G$  and  $H$ , can be defined as follows:

$$I(G, H) = \sum_{i,j} p(g_i, h_j) \log_n \frac{p(g_i, h_j)}{p(g_i)p(h_j)}, \quad (11)$$

where  $p(g_i)$ ,  $p(h_j)$ , and  $p(g_i, h_j)$  are the individual and joint probability distributions of  $G$  and  $H$ , respectively.

For an agent  $a$ , two variables are used for computing its MI: its direction,  $d(a)$ , and its two-dimensional coordinates,  $(pos(a)_x, pos(a)_y)$ . We evenly divide the scene into bins, each of which has a size of  $R * R$ . In other words, the area of  $\text{bin}(i,j)$  is  $[(i * R, j * R), ((i + 1) * R, (j + 1) * R)]$ . Then, we assign all the agents to those bins according to their 2D coordinates. In this way, we can compute the marginal probabilities as follows.

$$p(x_i) = \frac{\sum_{a \in C} pos(a)_x \in [i * R, (i + 1) * R]}{N}, \quad (12)$$

$$p(y_j) = \frac{\sum_{a \in C} pos(a)_y \in [j * R, (j + 1) * R]}{N}, \quad (13)$$

where  $N$  is the total number of agents in the simulation. Similarly, we can obtain  $p(X, Y)$ .

For the direction data, we process it in a similar way. Basically, we evenly divide the directional angles of a full circumference into  $K$  sectors; for any sector  $i$ , its covered direction range is from  $i * 2\pi/K$  to  $(i + 1) * 2\pi/K$ . So, the probability can be computed as follows:

$$p(d_i) = \frac{\sum_{a \in C} d(a) \in \text{sector}_i}{N}, \quad (14)$$

where  $N$  is the total number of agents in the simulation. In a similar manner,  $p(X, D)$  and  $p(Y, D)$  can be computed.

Based on the above Eq. 11, the MI between two distributions (i.e., the position distribution and the direction distribution),  $I_d$ , can be computed as follows.

$$I(X, D) = \sum_{i,j} p(x_i, d_j) \log_2 \frac{p(x_i, d_j)}{p(x_i)p(d_j)}. \quad (15)$$

$$I(Y, D) = \sum_{i,j} p(y_i, d_j) \log_2 \frac{p(y_i, d_j)}{p(y_i)p(d_j)}. \quad (16)$$

$$I_d = \frac{I(X, D) + I(Y, D)}{2}. \quad (17)$$

In a similar way, we can compute the MI between the agent position distribution and the agent velocity distribution,  $I_v$ . Since the positions of agents are well-dispersed, and the position distribution is balanced in general. But for the direction (or velocity) data of the agents, a better effect corresponds to harmonic heading (velocity), in other words, the corresponding direction (or velocity) distribution is ill-distributed; so the mutual information between the two distributions (i.e., the position distribution and the direction distribution) decreases when an aesthetic transformation appears. In this work, we adjust the movement of the agents based on the instant MI values,  $I_d$  and  $I_v$ . The main reason

of choosing the MI feedback in our approach is, based on our experiments, we found that the MI based measure is somewhat correlated with the fluency and stability of agent subgroups's localized movements in a crowd (refer to Section 5.3).

In the original SFM (Eq. 8),  $e_a$  represents the desired direction which is predefined with the assignment. Actually, it should be sensitive to the real-time environment. Assuming the agent  $a$  and its neighbors  $a'$  share the same bin  $bin(i, j)$ , we can redefine  $e_a$  with the MI feedback as follows.

$$e_a = e_a + \left( \sum_{a' \in bin(i, j), a' \neq a} \frac{|e_a - e_{a'}|}{num(i, j)} \right) \times I_d, \quad (18)$$

where  $num(i, j)$  is the number of agents in  $bin(i, j)$ . In this equation, the summing operator represents the average divergence of the agent  $a$  and its neighbors. After the MI feedback based adjustment, the direction for the agent  $a$  will keep its step with its neighbors.

Similarly, The desired velocity  $v_a^0$  in Eq. 8 can be adjusted according to  $I_v$  as follows.

$$v_a^0 = v_a^0 + \left( \sum_{a' \in bin(i, j), a' \neq a} \frac{|v_a^0 - v_{a'}^0|}{num(i, j)} \right) \times I_v \quad (19)$$

Finally, with the MI feedback to the original SFM model, we can control the agents' movement precisely and smoothly, which leads to a better visual effect of crowd formation transform.

### 3.4. Subgroup Clustering

Inspired by the work of [KLLT08, TYK\*09], we cluster individual agents in a crowd to subgroups to maximally maintain the formation of the collective sub-groups during the formation transition. In this work, we apply the Affinity Propagation clustering algorithm (AP) [FD07] to generate subgroups. The AP algorithm identifies a set of exemplars to best represent agents' positions in the formation. We choose the AP algorithm to cluster agents since an exemplar can be conceptually considered to represent the overall movement of its corresponding agent-subgroup, and the cluster number is determined adaptively and automatically. The details of the AP algorithm can be found at [FD07]. Note that we only apply the AP algorithm to the position set of the source formation.

The core of the AP algorithm is the design of a similarity measure. In this work, we design a novel and effective similarity measure using the local relative distance variance for the collective subgroup clustering. Thus, the source formation of individual agents can be represented as a weighted undirected complete graph  $G^*(S, E)$ , where the vertices  $S$  represents the individual agents' positions,

and  $E$  contains all possible edges that connect two distinct vertices. The edge between  $s_i$  and  $s_j$  in  $G^*$  is denoted as  $e(s_i, s_j)$ , and the weight for the edge  $e(s_i, s_j)$ , denoted as  $w(s_i, s_j)$ , is measured using the Manhattan distance. As described in Section 3.2, for  $\forall s_i \in S$ , its target position  $t_k$  is determined as  $t_k = \Psi(s_i) \in T$ . For each edge  $e(s_i, s_j)$  in  $G^*$ , its corresponding edge and weight in the target formation are  $e(\Psi(s_i), \Psi(s_j))$  and  $w(\Psi(s_i), \Psi(s_j))$ , respectively. Then, the similarity measure for subgroup clustering is defined as the square of the relative distance variance  $\Delta dis$  between each edge  $e(s_i, s_j)$  in the source formation and its corresponding edge  $e(\Psi(s_i), \Psi(s_j))$  in the target formation, detailed below.

$$\Delta dis(s_i, s_j) = |w(s_i, s_j) - w(\Psi(s_i), \Psi(s_j))|^2 \quad (20)$$

Finally, we can obtain the clustered collective subgroups in the source formation, which is a partition of the position set  $S$  of the source formation. Let  $Sub(S) = \{group_k\}_{k=1}^M$  be this partition of  $S = \bigcup_{k=1}^M group_k$ , where  $M$  is the number of clustered subgroups, and  $\forall group_i, group_j \in Sub(S)$  ( $i \neq j$ )  $\implies group_i \cap group_j = \emptyset$ .

### 3.5. Movement Control with Subgroup Constraints

Recalling the contents in the above Section 3.3, we adapt the SFM model to control the agents' motion with mutual information. In this work, we further extend the adapted SFM with a subgroup formation constraint as follows. For agents  $a, b \in group_k^S$ , we employ an additional attractiveness force,  $f_{sf\_attractive}$ , and for  $a \in group_k^S, b \notin group_k^S$ , we add an additional repulse force,  $f_{sf\_repulse}$ , as follows:

$$f_{sf\_attractive}(pos(a), pos(b)) = A' e^{|pos(a), pos(b)|}, \quad (21)$$

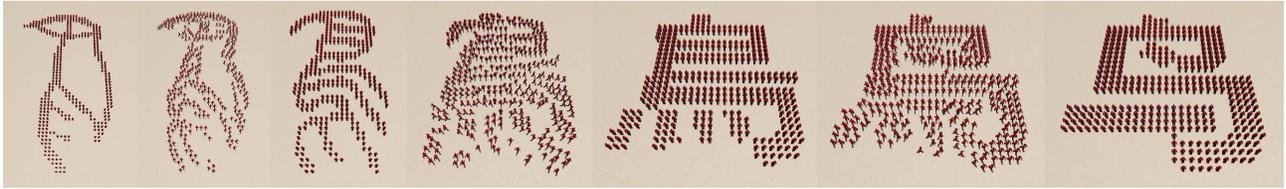
$$f_{sf\_repulse}(pos(a), pos(b)) = \frac{A' e^{-\frac{|pos(a), pos(b)|}{B'}}}{|pos(a), pos(b)|}, \quad (22)$$

where  $A'$  and  $B'$  are constants, and  $pos(a), pos(b)$  have the same definitions as in Eqs. 8-10.

Finally, the resultant force  $F$  is updated using the following form:

$$\begin{aligned} F = & f_{drive}(v_a) + f_{obstacle}(pos(o)) + \sum f_{repulse}(pos(a), pos(b)) \\ & + \sum_{a, b \in group_k} f_{sf\_attractive}(pos(a), pos(b)) \\ & + \sum_{a \in group_k, b \notin group_k} f_{sf\_repulse}(pos(a), pos(b)) + \xi_a(a \neq b). \end{aligned} \quad (23)$$

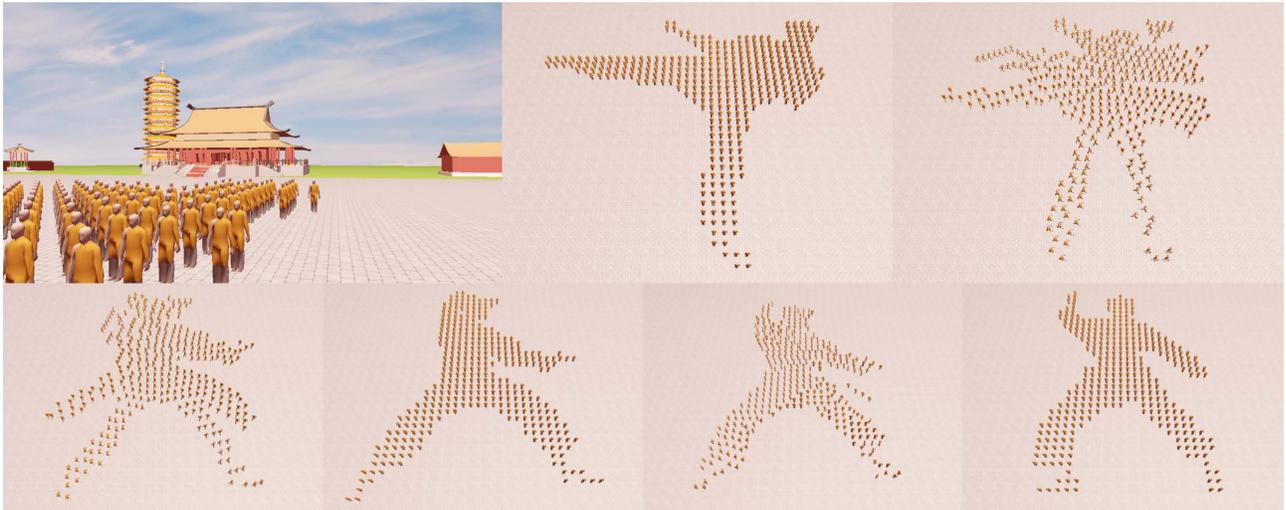
The added subgroup forces can help to maintain the local stability of all the subgroups during the formation transform process.



**Figure 2:** The evolution of the Chinese word "bird" from its oracle form to its modern simplified form. About 400 agents are used in this crowd simulation.



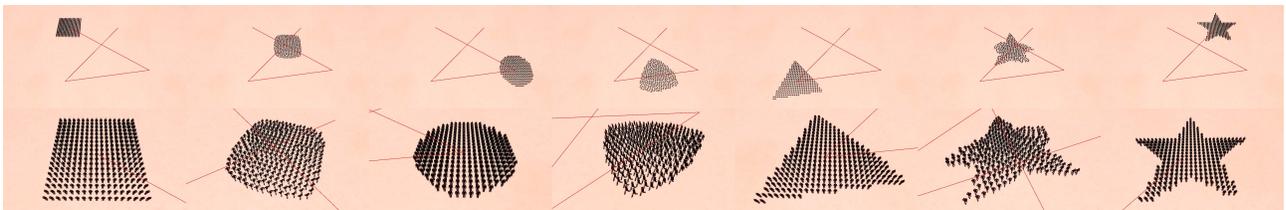
**Figure 3:** Crowd simulation of a "running" formation in a performance show (about 600 agents).



**Figure 4:** A group of 400 virtual monks (agents) presents the martial arts before a temple.



**Figure 5:** *Fighters are trained before a city gate in ancient time. The formations are noted as “awl”, “goose”, “hook”, and “round”, respectively.*



**Figure 6:** *Experimental result of formation transform while following a path: the aerial view (top) and the close-up view (bottom).*



**Figure 7:** *Formation transform from ape to human while following a curve path: the overview of the transformation (top) and the close-up view (bottom).*



**Figure 8:** An example of crowd formation transformations with obstacle avoidance



**Figure 9:** An example of crowd formation transformation by our approach to handle collision avoidance between two agent groups that pass through each other

#### 4. Results and Performance

We implemented our approach with C++. All our experiments were performed on an off-the-shelf computer with Intel Core i7-3770 3.40G Hz CPU. Our program was developed on top of the Horde3D engine [Hor07]. We simulated various scenarios of crowd formation transforms using our approach. Some selected simulation scenarios include evolution of the Chinese word “fish” (Figure 2), stadium performance (Figure 3), martial arts (Figure 4), fighter training formations (Figure 5), and formation transforms while following a specified path (Figures 6 and 7). For animation results, please refer to our supplemental demo video.

We also experimented our approach with collision avoidances. As shown in Figure 8, when a static obstacle appears in the environment, our approach cannot handle it well, because the initial pair assignment step is solved by only considering the direct distances between the matched agents, and the obstacle factor is not taken into consideration in this step. However, as shown in this figure, when the obstacle is passed, the crowd formation is quickly assembled and make a smooth transition to the target formation. Figure 9 shows a more challenging case of handing collision avoidances between two groups of agents that pass through each other. In this case, when the two groups interact each other, both the formations cannot be maintained. We argue the main reason is the lacking of an appropriate collision avoidance strategy for subgroups.

##### 4.1. Performance

To analyze the runtime performance of our approach, we recorded the frames per second (FPS) and the memory usage for the above simulation scenarios (Table 1). As shown in Table 1, our method can support real-time simulation of

**Table 1:** The performance and memory usage statistics of our method

Number of Agents	110	306	783	1190	1936
FPS	295	104	40	27	17
Memory(MB)	73	66	71	86	96

a crowd formation transformation consisting of more than one thousand agents on an off-the-shelf computer. In our approach, the most computationally intensive part is probably the pair assignment step, which has the time complexity of  $O(N^3)$ . However, the pair assignment step is an off-line process, and the memory usage is still affordable for off-the-shelf computers. Therefore, our method can be conveniently integrated into many real-time graphics applications such as video games.

#### 5. Evaluations and Comparisons

##### 5.1. Objective Aesthetic Measures

To quantify the aesthetic aspect of a crowd formation transformation, we introduce three different objective measures, as described below.

(1) Mutual information. Recall in Section 3.3, we describe how the mutual information concept can be adapted to measure the aesthetic aspect of a crowd formation transform, as well as how to compute the mutual information of a dynamic crowd. Specifically, in the result part, we use  $I_d$  (described in Eq. 17) as one of the objective measures to evaluate the aesthetic aspect of a crowd formation transform.

(2) The stability of local structure. The standard deviation and the average value can be used to quantify the stability property of local structure during a crowd formation

transformation. For an agent  $a_i$ , let  $min\_dis(i)$  be its minimum neighbor-distance (i.e., the minimum distance from it to its neighbors), and  $\mu_s$  is the average value of the minimum neighbor-distances for all the agents,  $(\sum_i min\_dis(i))/N$ . Then, the standard deviation can be computed as follows:

$$\sigma_{min\_dis} = \sqrt{\frac{1}{N} \sum_{i=1}^N (min\_dis(i) - \mu_s)^2}. \quad (24)$$

Clearly, the lower value of  $\sigma_{min\_dis}$  indicates that the agents have more similar distances from their neighbors; and vice versa.

(3) Effort balancing. We also employ the standard deviation and the average value to estimate the balancing of the agents' efforts. When a formation transformation is smooth and visually pleasing, for any agent, we anticipate that the effort from its source position to its current position is not only the least but also balanced.

For an agent  $a_i$ , its current effort is defined as follows:

$$curr\_eft(i) = dis(s_i, pos(a_i)), i = 1..N. \quad (25)$$

Here  $s_i$  and  $pos(a_i)$  denote its source position and its current position, respectively, and the function  $dis()$  computes the Euclidean distance. The standard deviation is calculated as follows:

$$\sigma_{curr\_eft} = \sqrt{\frac{1}{N} \sum_{i=1}^N (curr\_eft(i) - \mu_e)^2}. \quad (26)$$

Here  $\mu_e$  is the average value of all the agents' efforts,  $(\sum_{i=1}^N curr\_eft(i))/N$ .

## 5.2. Selection of the Optimal Distance Measure in Pair Assignment

In this work, we introduce a least effort distance measure (Eq. 3) in the pair assignment step (Section 3.2). In fact, there are three distance measure options for consideration: Euclidian distance, Manhattan distance, and our introduced least effort distance.

Figure 10 shows result comparisons of using the above three different distance measures at the pair assignment step in our approach. Figure 11 shows the quantitative comparisons of the objective aesthetic measures (defined in Section 5.1) if the three different distance measures are used in the pair assignment step. Table 2 shows the comparison of the obtained average values when the three distance measures are used. As clearly illustrated in Figure 11 and Table 2, our introduced least effort distance measure leads to lowest mutual information, best stability of local structure, and most balanced among the three distance measure options. In other words, the least effort distance measure balances both the macroscopic and microscopic movements of the crowd transformation, and thus it can lead to the best visual effect among the three.

**Table 2:** Comparison of the average values when different distance measures are used in the pair assignment step

	time(sec)	Average Value		
		$I_d$	Minimum	Effort
Least-effort	23.82	0.131	65.347	13.269
Euclidian	78.072	0.297	148.479	108.603
Manhattan	133.034	0.568	120.482	131.216

## 5.3. Comparisons with/without the Mutual Information Feedback

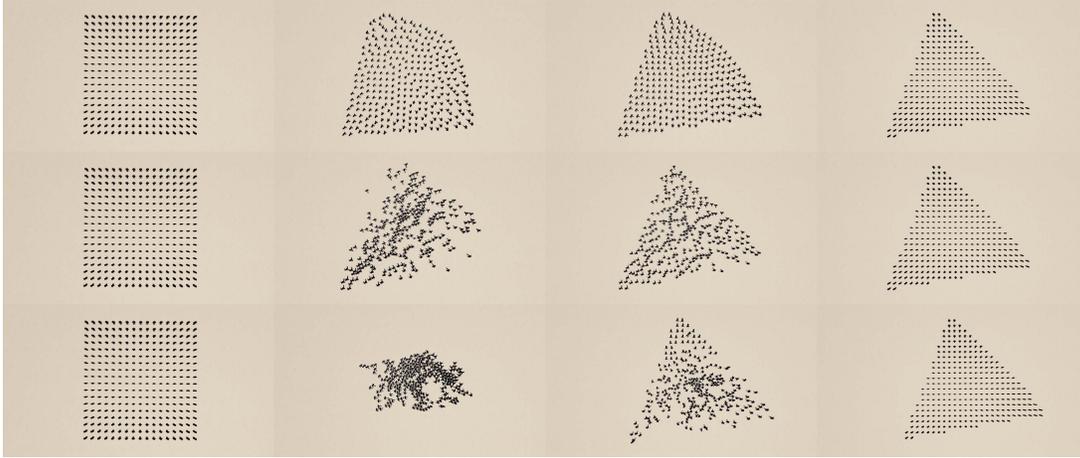
We also performed result comparisons with and without the MI feedback. One such example is shown in Figure 12. As shown in this figure, without the MI feedback, due to the influence of the repulse forces, the movement appear to have fluctuation during the transformation. In contrast, the introduced MI feedback can help to eliminate such a fluctuation phenomenon and thus improve the overall aesthetic effect of crowd formation transformation.

## 5.4. Comparisons with Selected State of the Art Methods

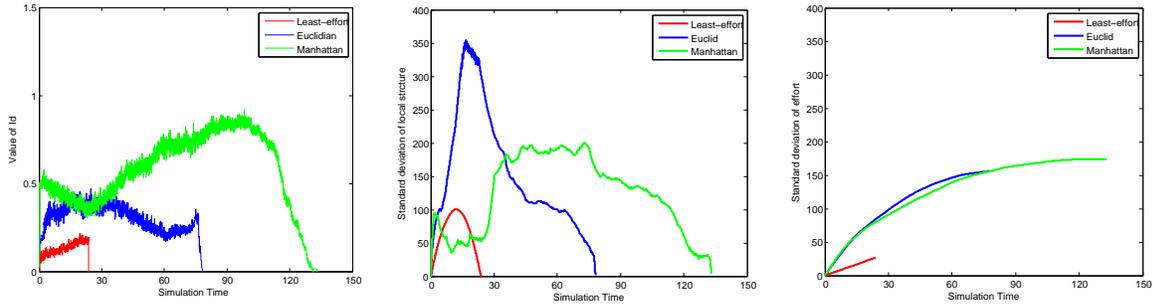
To evaluate our approach, we compared it with three widely-known crowd simulation approaches including RVO [Vd-BLM08], ORCA [VDBGLM11] and OpenSteer [Ope04]. RVO is a widely used velocity-based crowd simulation model. ORCA is the latest model derived from the RVO, and it is very efficient for robot movement control. OpenSteer is an open-source program for constructing steering behaviors for autonomous characters based on the well-known Reynolds's flocking model [Rey87]. As shown in Figure 14, the four different methods (our approach, RVO, ORCA, and OpenSteer) are applied to the same source and target formations. The visual results show that our approach can achieve the best visual effect among the four approaches.

Comparisons of the same example's objective aesthetic measures are shown in Fig. 13. Its left panel shows that the resulting mutual information of the four approaches is very close. The main reason is that the same pair assignment scheme is used in the four approaches. However, as shown in the middle of Fig. 13, during the transformation the agents in our approach have most similar distances from their nearest neighbors, which indicates that the local structure in our approach is more stable than the other three approaches. In addition, as shown in the right of Fig. 13, the slope of our method's curve is more stable than the other three approaches, which means the real-time efforts of the agents in our approach are well proportioned, which is one of the key factors leading to the best visual effect among the four approaches.

The statistics data (including simulation time and the obtained average values) of the four approaches are also shown



**Figure 10:** A formation transform example of using different distance measures at the pair assignment step in our approach: the least effort distance (top), the Euclidian distance (middle), and the Manhattan distance (bottom).



**Figure 11:** Comparisons of the mutual information,  $I_d$  (left), the standard deviation of the stability of local structure,  $\sigma_{min\_dis}$  (middle), and the standard deviation of the agents' efforts,  $\sigma_{curr\_eft}$  (right), if different distance measures are used in the pair assignment step.

**Table 3:** Statistics comparison of four different crowd simulation approaches (Our approach, ORCA, RVO and OpenSteer)

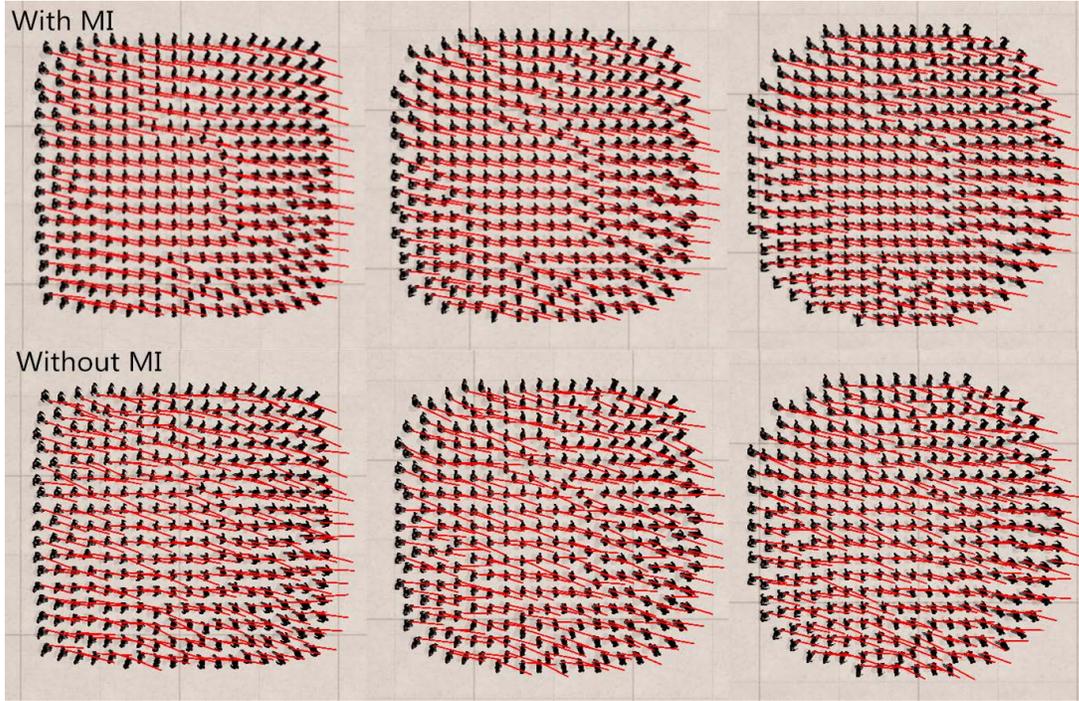
	time(sec)	Average Value		
		$I_d$	Minimum	Effort
our method	28.192	0.497	18.49	197.416
ORCA	32.159	0.519	18.636	239.07
RVO	35.892	0.56	18.735	238.575
OpenSteer	38.008	0.617	19.716	241.494

in Table 3. From this table, we can observe that our method has the least simulation time and achieves the best transformation among the four approaches, with respect to our three defined objective aesthetic measures.

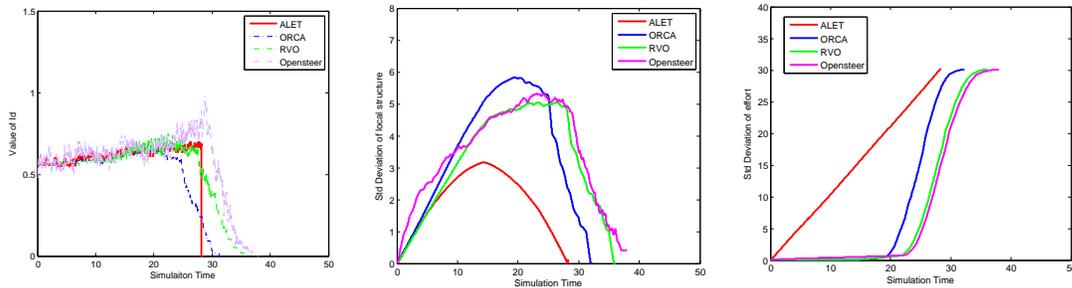
## 6. Discussion and Conclusions

In this paper, we introduce a novel crowd formation transform approach based on the principle of least effort, which helps to lead to visually pleasing formation transformations. The main idea of our approach is to achieve the aesthetic effect of crowd formation transform with least-effort pair assignment and a subgroup enhanced social force model. Our approach is able to effectively maintain the stability of local structure and preserve the dynamic collective behavior of a crowd during its formation transformation process.

Through many simulation experiments, we demonstrate that our approach is effective, robust and flexible for transforming crowd formations with various shapes and scales. In addition, through qualitative and quantitative comparisons between our approach and existing crowd simulation approaches, we show our approach can produce more visually



**Figure 12:** Result comparisons of our approach with and without the MI feedback. As shown in the figure, the result with the MI feedback is more stable with respect to the moving directions of the agents.

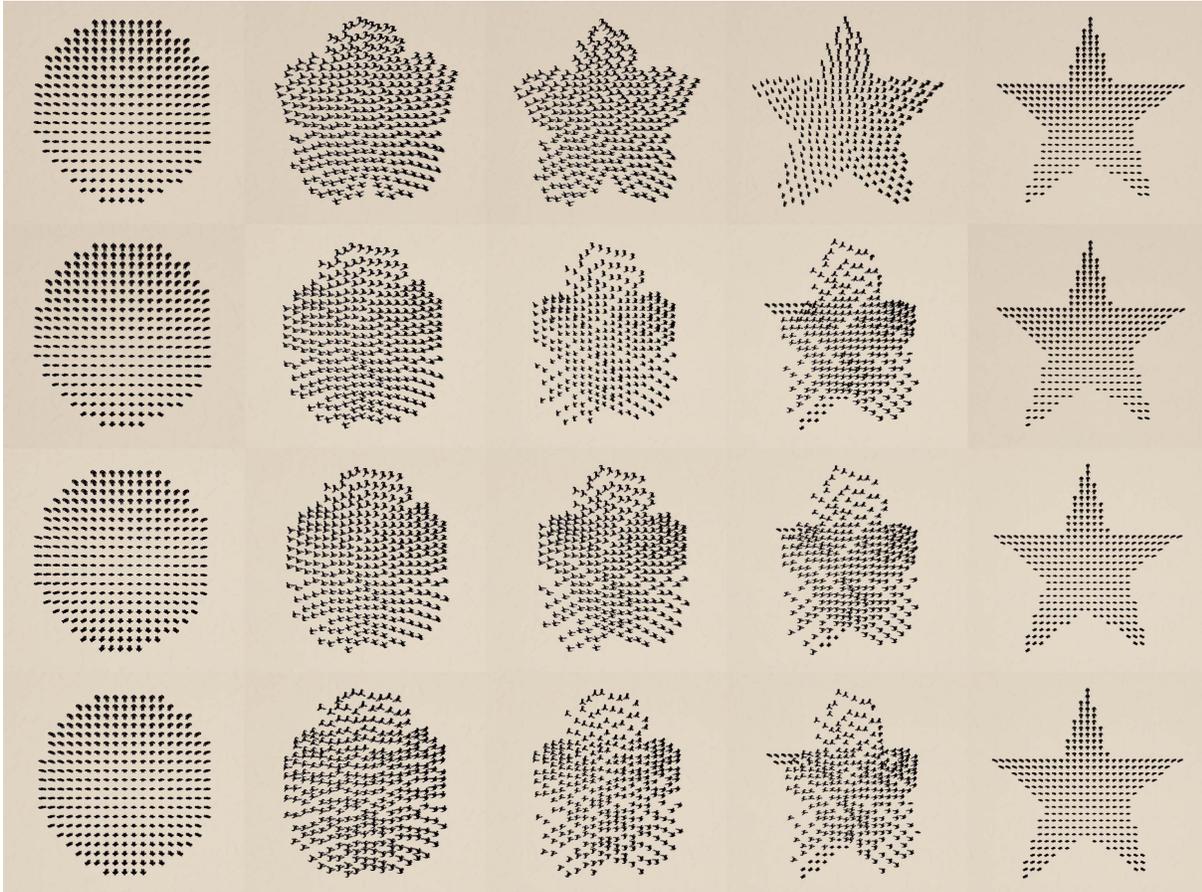


**Figure 13:** Comparisons of the mutual information,  $I_d$  (left), the standard deviation for the agents' local structure,  $\sigma_{min\_dis}$  (middle), and the standard deviation of the agents' efforts,  $\sigma_{curr\_eft}$  (right), if four different approaches (Our approach, ORCA, RVO and OpenSteer) are applied to the same source and target formations.

pleasing formation transformations than the three selected state-of-the-art approaches.

Certain limitations still exist in our current approach. First, the current work only focuses on geometric relationship among agents in a crowd, and many other group behaviors (e.g., behavior of competing cohorts crossing paths [GNL13]) cannot be handled well. Indeed, geometric relationship is just one of many collective behaviors exhibited by a crowd. In practice, many other factors and relations among agents also come to play and jointly affect the forming of

subgroups. In the future, we plan to investigate novel ways to further enrich the stability of agent subgroups. Second, our current approach cannot maintain formations when avoiding obstacles or other agents in the environment. We plan to investigate how to extend the current approach beyond immutable agents, for example, smoothly handling obstacles while maximally preserving the crowd formation, dealing with the mobility and size diversity with variations over time, as well as simulating more plausible sub-group transitions during the crowd transform process.



**Figure 14:** Comparisons of four different approaches given the same source and target formations (from top to bottom): Our approach, ORCA, RVO, and OpenSteer.

In the future, we plan to extend our current approach from 2D to 3D, and thus we can use it to simulate various formation transforms for certain crowds such as insect crowds. In addition, we also plan to explore whether and how the objective crowd formation measures introduced in this work can be used to quantitatively evaluate the deviation between the simulated crowd and the desired (target) formation.

#### Acknowledgments

This work was in part supported by the Natural Science Foundation of China (NSFC) (grants 61202207, 61100086, and 61170214), the Joint Research Fund for Overseas Chinese, Hong Kong and Macao Young Scientists of the National Natural Science Foundation of China (Grant No. 61328204), the Zhejiang Province Natural Science Fund for Distinguished Young Scientists (Grant LR12F02001), the China Postdoctoral Science Foundation (grant 2012M520067), and the Open Project Program of the State Key Lab of CAD&CG at Zhejiang University (grant

A1209). We also would like to thank the anonymous CGF reviewers for their constructive comments.

#### References

- [BB10] BRIDSON R., BATTY C.: Computational physics in film. *Science* 330, 6012 (2010), 1756–1757. [1](#)
- [FD07] FREY B. J., DUECK D.: Clustering by passing messages between data points. *Science* 315, 5814 (2007), 972–976. [3](#), [6](#)
- [FH85] FLASH T., HOGAN N.: The coordination of arm movements: an experimentally confirmed mathematical model. *The journal of Neuroscience* 5, 7 (1985), 1688–1703. [1](#)
- [FR12] FLAGG M., REHG J.: Video-based crowd synthesis. [1](#)
- [GCC\*10] GUY S., CHHUGANI J., CURTIS S., DUBEY P., LIN M., MANOCHA D.: Pedestrians: a least-effort approach to crowd simulation. In *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (2010), Eurographics Association, pp. 119–128. [1](#), [2](#)
- [GCLM12] GUY S., CURTIS S., LIN M., MANOCHA D.: Least-effort trajectories lead to emergent crowd behaviors. *Physical Review E* 85, 1 (2012), 016110. [2](#)

- [GD11a] GU Q., DENG Z.: Context-aware motion diversification for crowd simulation. *IEEE Computer Graphics and Applications* (Sept. 2011), 14–25. 2
- [GD11b] GU Q., DENG Z.: Formation sketching: an approach to stylize groups in crowd simulation. In *Proceedings of Graphics Interface 2011* (2011), Canadian Human-Computer Communications Society, pp. 1–8. 1, 2, 3
- [GD13] GU Q., DENG Z.: Generating freestyle group formation in agent-based crowd simulation. *IEEE Computer Graphics and Applications* (Jan. 2013), 14–25. 1, 2, 3
- [GKLM11] GUY S., KIM S., LIN M., MANOCHA D.: Simulating heterogeneous crowd behaviors using personality trait theory. In *Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (2011), ACM, pp. 43–52. 2
- [GNL13] GOLAS A., NARAIN R., LIN M.: Hybrid long-range collision avoidance for crowd simulation. In *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games* (2013), ACM, pp. 29–36. 2, 12
- [GSA\*09] GAYLE R., SUD A., ANDERSEN E., GUY S. J., LIN M. C., MANOCHA D.: Interactive navigation of heterogeneous agents using adaptive roadmaps. *IEEE Transactions on Visualization and Computer Graphics* 15, 1 (Jan. 2009), 34–48. 2
- [HFV00] HELBING D., FARKAS I., VICSEK T.: Simulating dynamical features of escape panic. *Nature* 407, 6803 (2000), 487–490. 2, 3, 5
- [HM95] HELBING D., MOLNAR P.: Social force model for pedestrian dynamics. *Physical review E* 51, 5 (1995), 4282. 2, 3, 5
- [Hor07] HORDE3D: [www.horde3d.org](http://www.horde3d.org). 9
- [HSK12] HENRY J., SHUM H., KOMURA T.: Environment-aware real-time crowd control. In *Eurographics/ACM SIGGRAPH Symposium on Computer Animation* (2012), The Eurographics Association, pp. 193–200. 1, 2, 3
- [HSK14] HENRY J., SHUM H., KOMURA T.: Interactive formation control in complex environments. *IEEE Transactions on Visualization and Computer Graphics* 20, 2 (Feb 2014), 211–222. 1, 2, 3
- [JCP\*10] JU E., CHOI M. G., PARK M., LEE J., LEE K. H., TAKAHASHI S.: Morphable crowds. *ACM Trans. Graph.* 29 (2010), 140. 2
- [KLLT08] KWON T., LEE K., LEE J., TAKAHASHI S.: Group motion editing. In *ACM Transactions on Graphics (TOG)* (2008), vol. 27, ACM, p. 80. 1, 2, 6
- [KOOP11] KULPA R., OLIVIERXS A., ONDŘEJ J., PETTRÉ J.: Imperceptible relaxation of collision avoidance constraints in virtual crowds. *ACM Transactions on Graphics (TOG)* 30, 6 (2011), 138. 2
- [KT11] KLOTSMAN M., TAL A.: Animation of flocks flying in line formations. *Artificial Life* 18, 1 (2011), 91–105. 2
- [Kuh55] KUHN H. W.: The hungarian method for the assignment problem. *Naval research logistics quarterly* 2, 1-2 (1955), 83–97. 5
- [LCS\*12] LI Y., CHRISTIE M., SIRET O., KULPA R., PETTRÉ J.: Cloning crowd motions. In *Eurographics/ACM SIGGRAPH Symposium on Computer Animation* (2012), The Eurographics Association, pp. 201–210. 2
- [LJK\*12] LEMERCIER S., JELIC A., KULPA R., HUA J., FEHRENBACH J., DEGOND P., APPERT-ROLLAND C., DONIKIAN S., PETTRÉ J.: Realistic following behaviors for crowd simulation. In *Computer Graphics Forum* (2012), vol. 31, Wiley Online Library, pp. 489–498. 2
- [LKF05] LAKOBA T., KAUP D., FINKELSTEIN N.: Modifications of the helbing-molnar-farkas-vicsek social force model for pedestrian evolution. *Simulation* 81, 5 (2005), 339–352. 2
- [LS80] LEE D., SCHACHTER B.: Two algorithms for constructing a delaunay triangulation. *International Journal of Parallel Programming* 9, 3 (1980), 219–242. 3, 4
- [MPG\*10] MOUSSAÏD M., PEROZO N., GARNIER S., HELBING D., THERAULAZ G.: The walking behaviour of pedestrian social groups and its impact on crowd dynamics. *PLoS one* 5, 4 (2010), e10047. 3, 5
- [Mun57] MUNKRES J.: Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial & Applied Mathematics* 5, 1 (1957), 32–38. 5
- [Ope04] OPENSTEER: [www.sourceforge.net/projects/opensteer](http://www.sourceforge.net/projects/opensteer). 2, 10
- [OPOD10] ONDŘEJ J., PETTRÉ J., OLIVIER A., DONIKIAN S.: A synthetic-vision based steering approach for crowd simulation. *ACM Transactions on Graphics (TOG)* 29, 4 (2010), 123. 2
- [PAB07] PELECHANO N., ALLBECK J., BADLER N.: Controlling individual agents in high-density crowd simulation. In *Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation* (2007), Eurographics Association, pp. 99–108. 2
- [PvdBC\*11] PATIL S., VAN DEN BERG J. P., CURTIS S., LIN M. C., MANOCHA D.: Directing crowd simulations using navigation fields. 244–254. 2
- [Rey87] REYNOLDS C.: Flocks, herds and schools: A distributed behavioral model. In *ACM SIGGRAPH Computer Graphics* (1987), vol. 21, ACM, pp. 25–34. 2, 10
- [SAC\*08] SUD A., ANDERSEN E., CURTIS S., LIN M. C., MANOCHA D.: Real-time path planning in dynamic virtual environments using multiagent navigation graphs. *IEEE Transactions on Visualization and Computer Graphics* 14, 3 (May 2008), 526–538. 2
- [TCP06] TREUILLE A., COOPER S., POPOVIĆ Z.: Continuum crowds. In *ACM Transactions on Graphics (TOG)* (2006), vol. 25, ACM, pp. 1160–1168. 2
- [TKB11] TURKAY C., KOC E., BALCISOY S.: Integrating information theory in agent-based crowd simulation behavior models. *The Computer Journal* 54, 11 (2011), 1810–1820. 2
- [TYK\*09] TAKAHASHI S., YOSHIDA K., KWON T., LEE K., LEE J., SHIN S.: Spectral-based group formation control. In *Computer Graphics Forum* (2009), vol. 28, Wiley Online Library, pp. 639–648. 1, 2, 6
- [VDBGLM11] VAN DEN BERG J., GUY S., LIN M., MANOCHA D.: Reciprocal n-body collision avoidance. *Robotics Research* (2011), 3–19. 2, 10
- [VdBLM08] VAN DEN BERG J., LIN M., MANOCHA D.: Reciprocal velocity obstacles for real-time multi-agent navigation. In *IEEE International Conference on Robotics and Automation* (2008), IEEE, pp. 1928–1935. 2, 10
- [XWY12] XU M., WU Y., YE Y.: Smooth and efficient crowd transformation. In *Proceedings of the 20th ACM international conference on Multimedia* (2012), ACM, pp. 1189–1192. 3
- [ZCC\*10] ZHOU S., CHEN D., CAI W., LUO L., LOW M. Y. H., TIAN F., TAY V. S.-H., ONG D. W. S., HAMILTON B. D.: Crowd modeling and simulation technologies. *ACM Trans. Model. Comput. Simul.* 20, 4 (Nov. 2010), 20:1–20:35. 2
- [Zip49] ZIPF G.: *Human behavior and the principle of least effort*. addison-wesley press, 1949. 1