J Supercomput (2013) 63:675–690
DOI 10.1007/s11227-011-0675-4

Massively parallel Modelling & Simulation of large crowd with GPGPU

Dan Chen · Lizhe Wang · Mingwei Tian · Jian Tian · Shuaiting Wang · Congcong Bian · Xiaoli Li

Published online: 23 September 2011

Abstract In peacekeeping, domestic, or combat operations, unanticipated crowd confrontations can occur. As a highly dynamic social group, human crowd in confrontation is a fascinating phenomenon. This paper presents a novel method based on the concept of vector field to formulate the way in which external stimuli may affect the behaviours of individuals in a crowd. Furthermore, Modelling & Simulation (M&S) of large crowds at individual level has long been placed in the highly computation intensive world. This study adopts GPGPU to sustain massively parallel M&S of a confrontation operation involving a large crowd. This approach enables investigation of a crowd consisting of tens of thousands individuals whose size was prohibitively large for conventional M&S technique to support. Experimental results indicate that the approach is efficient in terms of both performance and energy consumption.

Keywords Crowd Modelling & Simulation · GPGPU · Agent · Information entropy

D. Chen
School of Computer Science, China University of Geosciences, Wuhan, 430074, China

D. Chen
School of Computer Science, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK

L. Wang (✉)
Pervasive Technology Institute, Indiana University, 2719 East 10th Street, Bloomington, IN 47408, USA
e-mail: lizhe.wang@gmail.com

M. Tian · S. Wang · C. Bian · X. Li
Institute of Electrical Engineering, Yanshan University, Qinhuangdao, 066004, China

J. Tian
School of Economics, Huazhong University of Science & Technology, Wuhan, 430074, China

Fig. 1 Stampede at a musical festival in Duisburg (obtained from Internet)

1 Introduction

Human crowd is a fascinating social phenomenon in nature. A crowd of people may show well-organized structure and become disordered animals at another point. Numerous incidents/accidents in connection with crowd have been recorded in human history (see Fig. 1). How to predict and control the behavior of a crowd upon various conditions/events is an intriguing question faced by many psychologists, sociologists, physicists, and computer scientists. It is also a major concern of many government agencies when dealing with crowds in confrontation.

A crowd is not simply a collection of individuals and usually exhibits highly complex dynamics. Study of crowd in confrontation operations has received more and more attention. Crowd gatherings accompanied by severe violence have occurred frequently in nowadays restless world. From a researchers' perspective, another important reason is that the dynamics a crowd in a confrontation operation is largely influenced by external stimuli (properties and status of entities in the scenario as well as events), which are highly uncertain and often interact with the collective behavior of the crowd.

In general, pure mathematical approaches or analytic models are not adequate in characterizing the dynamics of a crowd. Crowd modeling and simulation (M&S) has recently been gaining tremendous momentum [21]. Existing models are largely at two extreme levels (microscopic and macroscopic): either model each individual as an autonomous agent, or treat the crowd as a whole [6, 9, 10] or consisting of homogeneous particles [7] with no cognitive features.

No matter a crowd is formed spontaneously or organized, individuals in the crowd gathering at the same time and space will globally exhibit common features, which can be well described by macroscopic modeling approaches. But the inherent pitfall of macroscopic modeling approaches is the incapability to reflect the impact brought by regional events and individualities within the crowd: for what microscopic approaches are designed. On the other hand, in general, there is lack of a formal method to formulate a crowd's common features with agent-based approaches.

The collective behavior of a crowd in a confrontation operation is determined by both unanticipated external stimuli and the common features of the crowd itself. In this study, we propose a novel method based on the concept of vector field to formulate the way in which external stimuli may affect the tendency of the behaviors of individuals. Our approach represents each individual as an autonomous agent whose actions are guided by the vector field model. As such, we bridge gap between the macroscopic with the agent-based approaches to more accurately characterize the interaction dynamics between a crowd and external stimuli.

Furthermore, as spotted out by Helbing, "Emotions play a decisive role in how people behave in crowds" and "the more nervous crowds get, the more unpredictably and irrationally they behave." Existing crowd models normally incorporate a number of tangible factors (such as speed, location, appearance, age) and some also consider intangible (such as emotional) factors. How to properly portray intangible factors and to quantitatively measure the impact these factors in a model remain a research challenge. This study also explores an information entropy based method to quantify the degree of panic of individuals and proposes the potential of disorder of the whole crowd. A quantitative analysis on intangible dynamics of a crowd in confrontation is then enabled.

It is a research challenge to support confrontation operation simulations (COS) involving large crowd. Large agent-based systems, such as simulation of large crowds at individual level, have long been placed it in the highly computation intensive world. Using traditional CPU-based high performance computing technology may provide an ad hoc solution to the performance issue [15–20] but this type of technology is being subject to a number of limitations: heat dissemination, excessive energy consumption, high-density power, and excessive cost for associated cooling systems. There exists a pressing need for computing methods for COS which can simulate a crowd of a large size while ensuring energy efficiency.

In the past few years, the modern Graphics Processing Unit (GPU) has evolved into a highly parallel, multithreaded, and many-core processor far beyond a graphics engine which substantially outpaces its CPU counterparts in dealing with computationally demanding, complex problems [3]. In this study, we have developed a parallelized crowd simulation approach, which successfully adopts General-purpose computing on the graphics processing unit (GPGPU) to thoroughly exploit the parallelisms of the COS process. The proposed approach has been developed upon NVIDIA's Compute Unified Device Architecture (CUDA) [5], a general purpose parallel computing architecture. Results demonstrate that GPGPU-aided approaches are remarkably superior to the distributed computing based counterparts in terms of both performance and energy consumption.

A shorter version of this article was presented in the 16th IEEE International Conference on Parallel and Distributed Systems [1], which is mainly about the agent-based model of COS. The rest of the paper is organized as follows. Section 2 provides some background knowledge and redefines notations borrowed from other disciplines. Section 3 introduces the vector field approach. Section 4 presents a case study of simulation of a crowd in confrontation operation. This section also gives a quantitative analysis on the evolvement of the simulated crowd's entropy in terms of degree of panic. Section 5 introduces the approach to M&S of large crowd in confrontation operation using GPGPU. We conclude this paper with a summary and future work in Sect. 6.

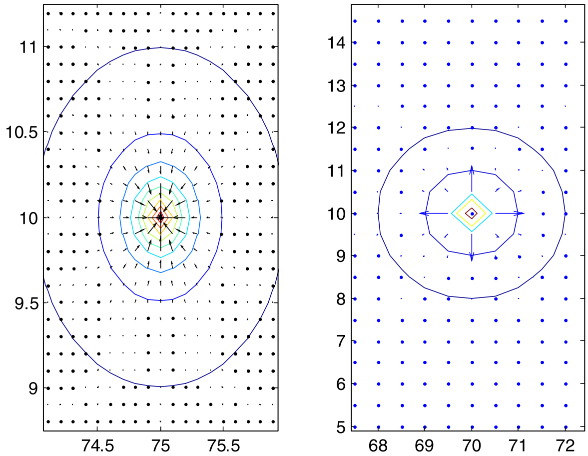


Fig. 2 A gravitational field (*left*) and a repulsive field (*right*)

2 Background and notations

This study (1) adopts agent-based approach for modeling individuals; (2) transplants the concept of vector field to reflect the influences of external stimuli on a crowd; (3) uses information entropy to analyze a crowd's intangible dynamics; (4) adopts GPGPU to parallelize COS to sustain scenarios consisting of large crowds. Several important notations related to the above methods are presented as follows:

- The *Agent-based approach* is currently the most active approach used for crowd M&S in the community of computer science and engineering [2–5, 16, 20]. A crowd is regarded as a collection of heterogeneous individuals that are empowered with decision-making capability, with each agent representing an individual. The agent-based approach is the most natural way to model behaviors with strong individual differentiations. A typical example is the PAX system which provides an M&S tool for scenarios of peace support missions [13]. Our model also uses autonomous agents to model individuals.
- A *vector field* is a construction in vector calculus which associates a vector to every point in a subset of Euclidean space. Vector fields are often used in physics to model the strength and direction of some force, such as the magnetic or gravitational force, as it changes from point to point. A vector field in the context of this study formulates the relationship between people's intention (internal) and the external stimuli. The vector field theory in physics will be adopted as the mathematic basis of the macroscopic model. A vector field (see Fig. 2) describes the influence ("force") that a stimulus has on a certain intention of a person who perceives this stimulus in the scenario's space. An intention is regarded as a *charged particle*, whose *charge* is subject to the intention's *magnitude* and may change with the evolvement of simulation. A vector field is *dynamic* if its intensity and direction may change over time, otherwise it is *fixed*.

In this study, *a vector field* is defined for *one particular stimulus* which merely works on *one particular intention*. As such, interference does not exist between any two vector fields. This definition significantly differs from those vector fields in physics. The effect of multiple vector fields on an individual is only exhibited by the combined force resulted from the forces these vector fields act on multiple intentions exclusively.

- "*Entropy*" has important physical implications as the degree of "disorder" of a physical system [8, 12, 14]. Information entropy is a measure of the uncertainty associated with a random variable (X), which is defined as

$$H(X) = H(P_1, P_2, \dots, P_n) = -\sum P(x_i) \log P(x_i) \tag{1}$$

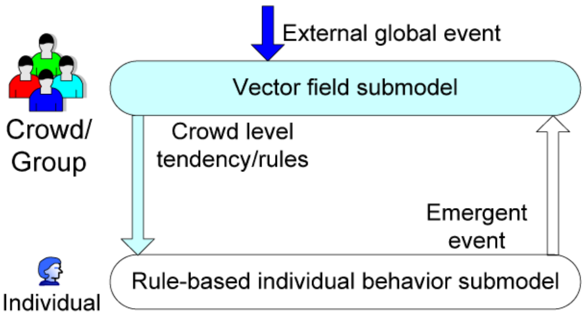
where $P(x_i)$ is the probability that X is in the state x_i, and $\sum_n P(x_i) = 1$. If $P(x_i) = 0$, $P(x_i) \log P(x_i)$ is defined as 0. The more disorderly a system is, the more information it contains, and vice versa.

In the context of our crowd model, five types of behaviors have been defined, namely "following," "avoiding," "adjusting," "confrontation," and "retreat." An agent may have specified a behavior at any point to conduct subject to a probability distribution on these candidate behaviors. The "information entropy" can then be calculated for the whole crowd. The value of the entropy will provide a quantitative measure on how disorderly the current crowd is, which may then facilitate controlling the crowd.
- The notation *Degree of Parallelism* (DoP) is referred to quantify the parallelism a problem to be solved. A problem's DoP means the number of portions in the problem, which can be concurrently solved/executed with the same results as those attained in a serial manner.

3 The hybrid behavior model

A hybrid behavior model is designed to manipulate each agent's behavior. Figure 3 presents a conceptual view of the proposed model consisting of two submodels, i.e., (1) a rule-based submodel to specify each agent's exact behavior and (2) a vector field submodel representing the influences of external stimuli common on all agents in the simulation. We adopted a classic design for an agent, which has the cognition capability to sense, deliberate then act. Rule-based agent approach has been extensively covered by existing work. This study emphasises the vector field submodel only.

Fig. 3 Conceptual view of the vector field submodel and the behavior model



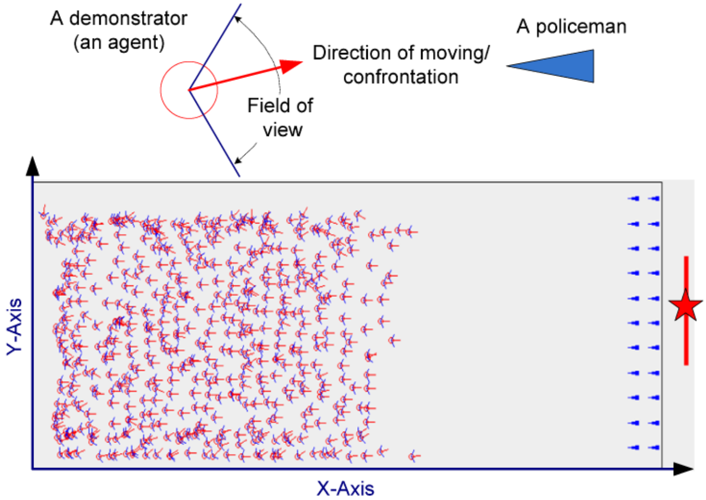


Fig. 4 Simulation of a crowd in a confrontation operation

The vector field submodel maintains a set of vector fields. The submodel computes the integrative influence of the external stimuli on people's various internal intentions, and it outputs the *tendency* (measured by the combined force) of an individual's behavior. The tendency means what the individual is likely to do rather than a *deterministic* action/motion as in Helbing's approach [7]. The vector field approach has been examined in a scenario of demonstration in front of a governmental building, as shown at the top of Fig. 4.

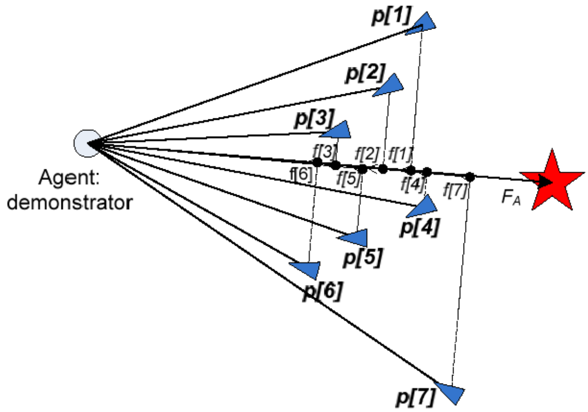
A crowd of demonstrators move on a march toward a governmental building (with its entrance highlighted as the *red star*). Each individual demonstrator is represented by a small red circle with its field of view and moving/confrontation direction marked. Armed policemen are denoted by blue triangles pointing at their moving/confrontation directions. Agents are moving within the 2D space confined by the upper and lower bounds.

Here, we consider an agent's intentions of two types: (A) "To approach the entrance of the governmental building," and (B) "To avoid being attacked by the policemen." A basic type of vector fields have been defined to represent the influences of the governmental building on an agent's intention A ($\vec{E}_g(\vec{r})$) in (2) and (2) each armed policeman on an agent's intention B ($\vec{E}_{p[i]}(\vec{r})$) in (4):

$$\vec{E}_g(\vec{r}) = \begin{cases} C_A \frac{(\vec{r}-R_g)}{|\vec{r}-R_g|} & |\vec{r} - R_g| < D_1 \\ k_A \frac{Q_g(\vec{r}-R_g)}{|\vec{r}-R_g|^3} & D_1 \le |\vec{r} - R_g| \le D_{\mathrm{MAX}} \\ 0 & |\vec{r} - R_g| \ge D_{\mathrm{MAX}} \end{cases} \tag{2}$$

$$\vec{E}_{p[i]}(\vec{r}) = \begin{cases} -k_B \frac{Q_{p[i]}(t)(\vec{r}-R_p)}{|\vec{r}-R_p|^3} & |\vec{r} - R_{P[i]}| < D_2 \\ 0 & |\vec{r} - R_{P[i]}| \ge D_3 \end{cases} \tag{3}$$

Fig. 5 Combined force acted on an agent



where Q_g is a fixed variable representing the "intensity" of the governmental building; $Q_p(t)$ represents the "intensity" of the governmental building; $\vec{r}$ is a vector from the origin of coordinate to an agent's location; R_G is the vector (fixed) from the origin to the star; R_G is the vector from the origin to a policeman (i); C_A is a constant. D_1, D_2, and $D_3(\gg D_2)$ are constants representing distances; D_{MAX} represents the max distance between any two locations in the scenario; k_{A} and k_{B} are two coefficients.

We write an agent's intentions A, B at a time point as $Q_{\text{A}}(t)$ and $Q_{\text{B}}(t)$, respectively. The combined effect of the governmental building and the policemen ($i = 1, 2, \ldots, n$) on the agent can be written as

$$\vec{F}_{\text{AB}}(t) = \vec{F}_{\text{A}}(t) + \vec{F}_{\text{B}}(t) = \vec{E}_g(\vec{r}) \times Q_A(t) + \sum_{i=1}^{n} \vec{E}_{p[i]}(\vec{r}) \times Q_B(t) \qquad (4)$$

In this scenario, policemen are deployed between the governmental building and the demonstrators, and it close to the government. Given that only intentions A and B are concerned, the agent tends to confront to policemen when $|\vec{F}_{\text{AB}}(t)|$ is small enough ($\leq \varepsilon$).

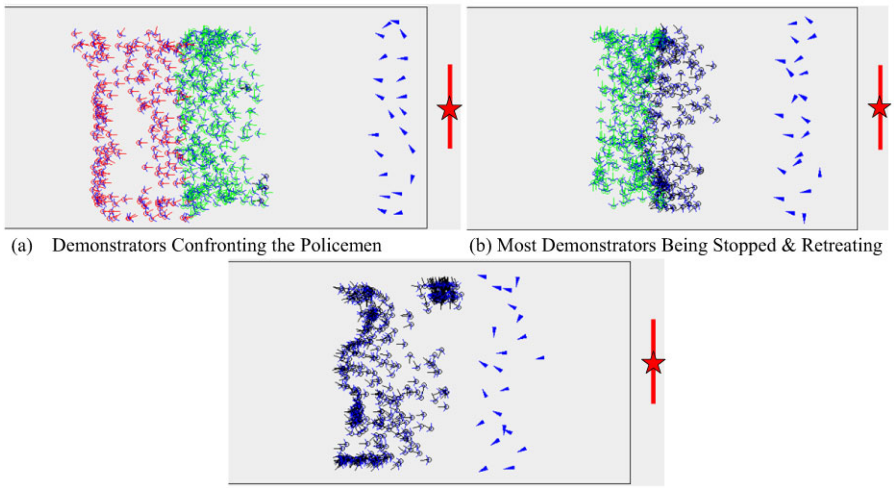
The effect of $\vec{F}_{\text{AB}}(t)$ on the agent relies on the component (A-component) of $\vec{F}_{\text{AB}}(t)$ along the direction of $\vec{F}_{\text{A}}(t)$, negative means in the same direction (see Fig. 5). Let α denote the angle between $\vec{F}_{\text{AB}}(t)$ and $\vec{F}_{\text{A}}(t)$, the magnitude of the A-component of $\vec{F}_{\text{AB}}(t)$, written as $\vec{F}_e(t)$:

$$|\vec{F}_e(t)| = |\vec{F}_{\text{AB}}(t)| \times \cos(\alpha) \qquad (5)$$

When $\vec{F}_e(t)$ is negative, the agent tends to leave the governmental building and policemen; otherwise, the agent tends to approach the governmental building. The intensity of the tendency of the agent's behavior is proportional to $|\vec{F}_e(t)|$.

4 A case study of confrontation operation simulation

Simulation has been executed using the agent model based on the vector-field method to examine the effectiveness of the proposed method. The dynamics of the simulated system has been quantified via entropy calculation afterward.



(a) Demonstrators Confronting the Policemen

(b) Most Demonstrators Being Stopped & Retreating

(c) Demonstrators Being Expelled

Fig. 6 Evolvement of a simulated crowd in a confrontation operation

4.1 Simulation of a crowd in a confrontation operation

The simulation scenario involves a crowd of 500 demonstrators marching to the governmental building and interacting with the policemen (22 on initialization) attempting to expel the demonstrators. The simulation lasts for 200 time units.

Figure 4 (see description in Sect. 3) illustrates the initial stage of the simulation. The gravity field, $\vec{E}_g(\vec{r})$, dominates the agents' (marked in red) tendency; $\vec{F}_e(t)$ on the agents that are positive, so the agents approach the red star. A small number of agents in the front lead the way followed by the rest of the agents.

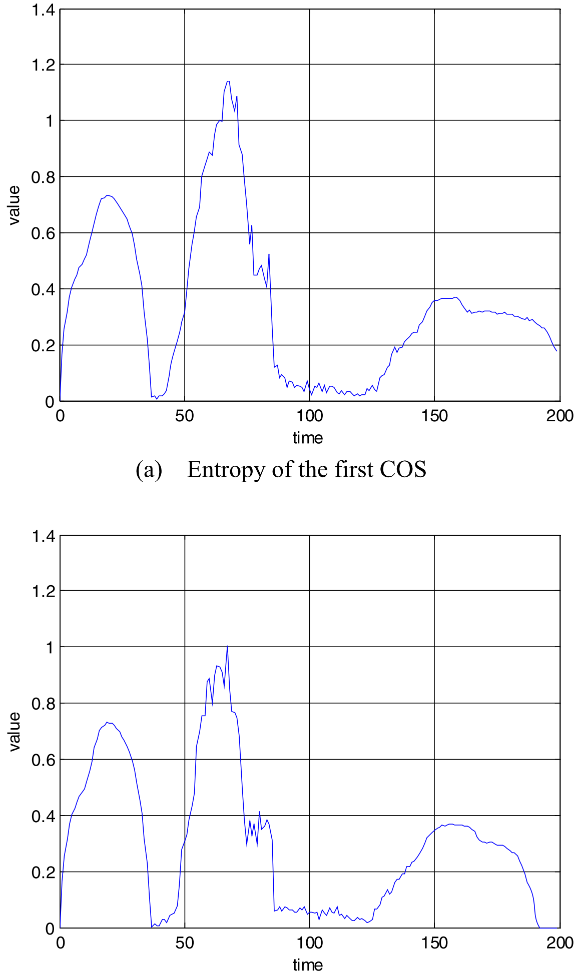
Figure 6 demonstrates three other stages of the simulation. When the agents in the front get close enough to the policemen, the repulsive fields, $\vec{E}_{p[i]}(1 \sim n)$, generate repulsive force great enough which balances $\vec{E}_g(\vec{r})$. $\vec{F}_e(t)$ on some agents in the front half become less than ε, thus most of them confront the policemen and their color turns to green (see Fig. 6(a)).

Policemen start to move toward the agents to prevent the demonstrator agents from further approaching the red star (see Fig. 6(b) and (c)). When the distance between the policemen and the agents in the front is generally small, the agents' panic level increases and the intensity of their intention A diminishes. In these cases, $\vec{E}_{p[i]}(1 \sim n)$ dominate, and $\vec{F}_e(t)$ of most agents become negative which drive more and more agents to leave the demonstration. During the stage illustrated by Fig. 6(d), there are always small numbers of agents that gather into groups. These groups will only dissolve and leave until the police come too close to them.

4.2 Dynamics analysis via entropy calculation

In this study, we introduce the concept of information entropy to analyze the degree of disorder of the simulated crowd. The definition of entropy is available in Sect. 2.

Fig. 7 The information entropy of the case studies of COS



(a) Entropy of the first COS

(b) Entropy of the second COS

Figure 7(a) presents the information entropy of the crowd calculated for the simulation scenario presented in Sect. 4.1:

- A peak in entropy in the duration from time 0 to 40 can be observed. This is caused by the diversity of agents' behaviors when the agents pour into the demonstration at the beginning.
- After time 20, the crowd marches to the governmental building and most agents' behaviors converge, thus the value of entropy drops sharply.
- From time 40 to 70, there is a sheer increase in information entropy. Agents are in the process of approaching the governmental building, and their degrees of panic are growing when they get closer to policemen.
- After time 70, agents continue to move forward and most of their behaviors become confrontation to the policemen. This results in a drop in the value of information entropy.

- From time 100 to 130, the policemen move toward the agents in advancing. Agents gradually start to “retreat.” As the number of retreating agents increases, the behaviors of the crowd tend to converge, and the entropy’s value drops further.
- When some small groups of agents are formed and remain in confrontation to the policemen, the third peak is reached at time 150. There can be a few agents who manage to cross the policemen and rush into the governmental building. This diversity makes the information entropy at a relatively high level.

In the previous simulation, agents with strong intention A may break through the policemen line. To test the influence of more policemen on the order of the crowd, we performed another simulation with more policemen (30) deployed when most agents are in the state of confrontation. From Fig. 7(b), we can see that the value of information entropy in the current simulation is generally less than that calculated from the previous simulation. The information entropy also drops to zero at the end stage of the simulation as a contrast to the results presented in Fig. 7(a). This denotes that the crowd is more orderly in this simulation and the probability of unanticipated emergency events is reduced. The time for the crowd to reach stability is also shortened. In the two simulations, the information entropy properly reflects the status and the evolvement of the dynamics of the crowd under the influence of the policemen and the governmental building.

5 Confrontation operation simulation aided by GPGPU

Due to the complexity of the COS at the individual agent level, the size of a simulated crowd is very limited, i.e., around 8,000 individuals to the maximum, even using a cutting-edge desktop computer. Another problem is the execution efficiency. We developed a GPGPU-aided solution to address these problems.

5.1 Parallelization of crowd simulation

The most intensive computation of a COS in this study lies in the execution of agents. The complexity increases almost linearly with the size of crowd, i.e., the number of agents. A sequential COS operates in a number of identical virtual time frames, typically representing 0.2 second in real world. For each time frame, the simulation executes all agents one by one to compute each agent’s velocity (V), position (P), decision on behavior selection (B), and its target (G). Computing the four attributes of an agent requires the COS’ system state (obtained from the last time frame), and this is independent from the results of any other agent at the current time frame.

This means that it is possible to parallelize the task of executing multiple agents. Considering the feature of the crowd simulation, we propose a scheme to partition the execution of agents into a number of subtasks with each executing an individual agent as shown in Fig. 8. Given a COS scenario consists of *NA* individuals, the COS’ DoP equals *NA* at the individual agent level.

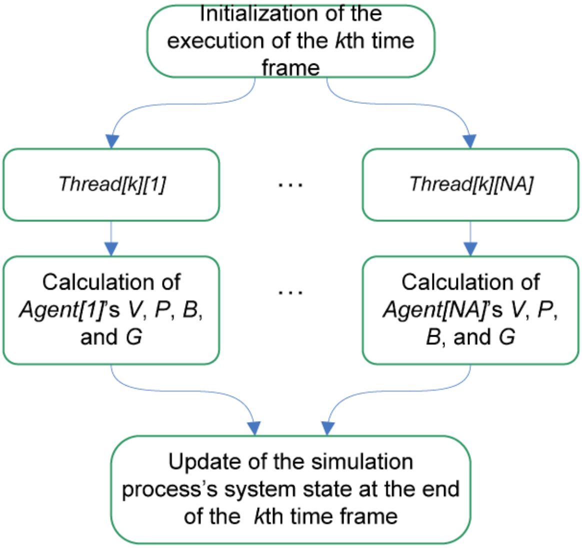


Fig. 8 Execution task graph of a parallelized simulation scheme

5.2 Evaluation of performance and energy efficiency

We have performed a number of experiments to study the performance of the alternative COS systems. The configurations of the test bed are presented in Table 1. The original sequential version of the COS has been modeled using Java upon *RePast 3.0* [2], a Java-based toolkit for the development of lightweight agents and agent models. RePast has become a popular and influential toolkit, providing the development platform for several large multiagent simulation experiments, particularly in the field of social phenomena. We first examined the overhead distribution of the simulation program. The computer node is only able to execute ∼8000 demonstrators to the maximum. Given a scenario consisting of 4,000 demonstrators with timestep set as 0.5 second (simulation time), the simulation execution time on a single computer node is ∼309 seconds. The elapse times in calculating V, P, B, and G is ∼308 seconds, which contributes about 99.6% of the overall overhead. Clearly, the performance bottleneck of the program lies with calculation of the four attributes of the agents.

5.2.1 GPGPU-aided confrontation operation simulation

Based on the above observations, we developed a parallelized simulation using GPGPU, which excels in handling a large number of concurrent fine-grained subtasks. The GPGPU-aided COS uses an individual CUDA thread to compute each agent's four attributes in each time frame (see Fig. 8). The new simulation program was developed based on JCuda (version 0.3.2a) [11], a Java binding for Java programs to interact with CUDA runtime and driver APIs. Thus, most of the original simulation code in Java can be reused while still having the benefit from the underlying powerful parallel programming and computing capabilities offered by GPGPU.

Table 1 Configurations of the testbed

Specifications	Computers	
	Desktop computer (Client)	Computer Cluster (Server, 1 master node and 15 worker nodes interlinked via 1 Gbps Ethernet)
Operating system	Windows 7 Professional	Rocks 4.2.1 (Cydonia) Red Hat Enterprise Linux ES release 4, X86_64 2.6.9-42.ELsmp
CPU	Intel Pentium Dual Core at 3.20 GHz and 3.19 GHz	2 × Intel Dual Core at 3.0 GHz
RAM	2048 M	4096 M
Power consumption	*650* W	*3.5* KW ∼ *14* KW
Specifications of NVIDIA GeForce GTX 480		
CUDA Cores	480	
Processor clock	1401 MHz	
Standard memory	1536 MB GDDR5	
Memory bandwidth	223.8 GB/sec	
Power consumption	*250* W	

Given that a COS scenario consists of *NA* individuals, we produced a scheme which maps this task to CUDA threads in the following steps:

Step 1: After initialization of the kth time frame, the host assigns *NA* data sets derived from the simulation's current system state. Each data set corresponds to an individual agent for computing its velocity, position, decision on behavior selection, and its target;

Step 2: The host invokes *NA* CUDA threads via JCuda, and these threads are evenly grouped into 480 blocks (*NA*/480 threads operating in each block) on the "device." This means 480 cores on the GPU are assigned to execute these agents with each core computing one agent's attributes individually;

Step 3: Step 2 repeats until the *NA* threads complete execution. The *NA* agents' new attributes are then passed from the device to the host. The host then updates system state through the RePast simulation engine and enters the $(k + 1)$th time frame.

Our design minimizes the thread number in a thread block while it creates thread blocks as many as possible when executing threads of this type. Hence, these threads can make the most of CUDA cores to deal with intensive computations and occupy as much fast shared memory (manipulated by each block) as possible to buffer the intermediate data generated during their executions. The GPGPU-aided COS can support scenarios consisting of more than 30,000 demonstrators.

Table 2 Performance of sequential and parallelized COS systems

Number of agents	Sequential COS	Cluster-aided COS (parallelized)		GPGPU-aided COS (massively parallelized)	
	Execution time (Sec)	Execution time (Sec)	Speedup	Execution time (Sec)	Speedup
3000	206	102	2	20	9.8
4000	309	138	2.2	31	10
5000	565	191	3	57	10
8000	1219	439	2.8	95	13
10000	N/A	680	N/A	115	N/A
12000	N/A	718	N/A	129	N/A
20000	N/A	1409	N/A	209	N/A
30000	N/A	N/A	N/A	317	N/A

5.2.2 Performance evaluation and energy efficiency analysis

In order to investigate the potentials of traditional CPU-based high performance techniques, e.g., cluster computing, and to establish a reference for evaluate the GPGPU-based approach, we developed another parallelized COS with the support of HLA_RePast [2], a middleware which supports the execution of multiple interacting instances of RePast agent-based models. Thus, a Cluster-aided COS (CCOS) has been established, and the load of executing the original COS can then be distributed over the 15 worker nodes of the computer cluster (see Table 1).

We performed a series of experiments which focuses on speedup (comparing to sequential COS, referred to as SCOS) and aims to investigate and compare the performance of GPGPU-aided and cluster-aided COS systems. Table 2 gives the execution times of the two types of COS systems with different numbers of agents (demonstrators). The results indicate that (1) the two parallelized COS systems significantly improve the runtime performance and scale well with the number of agents; (2) the GPGPU-aided COS (referred to as GCOS) always excels in performance improvement. The results (agent number $\leq$8000) are highlighted in Fig. 9.

The GPGPU-aided COS in this study operates on a graphic card with power consumption amounts to the maximum 250 W included in the maxim 650 W power consumption of the desktop. During the execution of GCOS on the desktop (viewed as a CPU-GPU hybrid system), the desktop's consumption was about $\sim$210 W. In contrast, during the execution of SCOS on the desktop (viewed as a pure CPU system), the desktop's consumption was about $\sim$130 W. The computer cluster's power consumption amounts to $\sim$9 KW during the execution of CCOS.

Taking the test scenario with 8,000 agents, for example, the execution times with SCOS, CCOS, and GCOS are 1219 s, 439 s, and 95 s, respectively. The total energy consumption using the three systems is $\sim$158470 J, $\sim$3951000 J, and $\sim$19950 J. Comparing to the GCOS system, SCOS/CCOS consumes $\sim$7/$\sim$197 times more energy. This analysis even does not consider the power consumption of the cooling system for the computer cluster room. The experimental results demonstrate the great

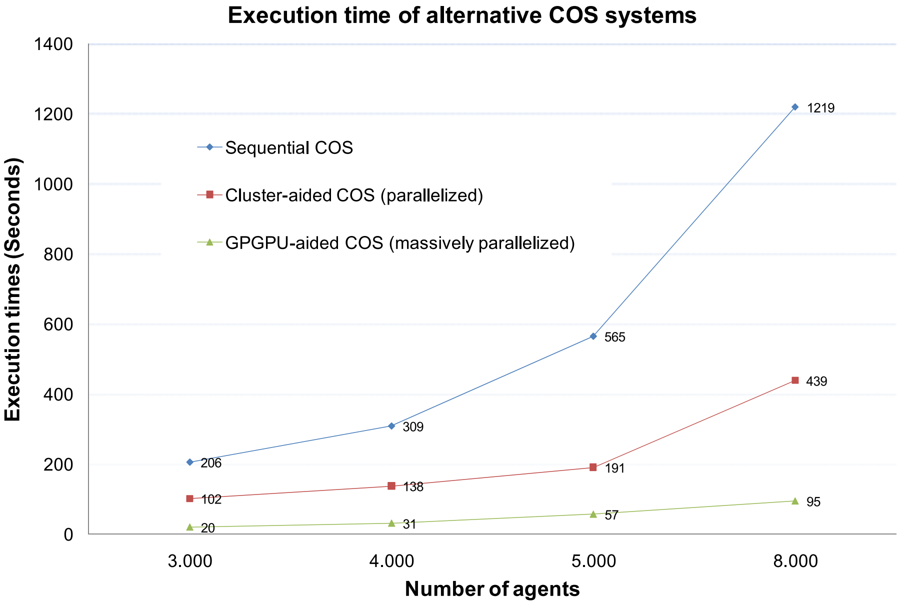


Fig. 9 Execution time of alternative COS systems (agent number ≤8000)

advantages of GCOS over SCOS and CCOS in terms of both runtime performance and energy consumption.

6 Conclusions and future work

This study explored an energy-efficient and high performance solution to simulation of confrontation operations involving large crowds. The novel simulation approach, namely GPGPU-aided COS, has been developed to address the scalability and performance issues using GPGPU.

We first proposed a vector field method which aims to formulate the way in which external stimuli may affect the tendency of the behaviors of individuals. Together with the agent-based approach, a model for simulation of crowd in confrontation operations has been established using RePast. We also introduced the concept of information entropy to analyze how the change of each individual's behavior may affect the intangible dynamics of the whole crowd. A case study of crowd simulation has been carried out. Through the measure of information entropy, the status and the evolvement of the dynamics of the crowd can be revealed. The results indicate that (1) the proposed COS model can exhibit typical behavior pattern of a crowd in confrontation; and (2) that information entropy can provide evident support to the design of control tactics for crowd control.

This study then emphasizes the feasibility and effectiveness of COS with GPGPU. The GPGPU-aided approach naturally divides a COS into a large number of fine-grained tasks, thus it effectively exploits the parallelism of the COS system at the individual agent level. It seamlessly maps the tasks to the same number of CUDA threads which can be executed concurrently by hundred of GPU cores.

Experiments have been carried out to evaluate the performance of GPU-aided COS and to investigate the potentials of traditional CPU-based high performance techniques. A cluster-aided COS has been developed upon HLA-RePast. Although cluster-aided COS runs over a high-end CPU-based computer cluster, GPU-aided COS prevails in runtime efficiency: G-EEMD is ∼6 times faster than the best distributed counterpart does. More importantly, the graphic card has maximum power consumption only $\sim 1/36$ of the computer cluster's power consumption. This figure does not consider the power consumption of the cooling system to ensure the computer cluster operable. The results indicate that GPGPU is a very promising technique in simulation of social phenomena. The proposed GPGPU-aided COS is indeed a highly energy-efficient and an ultra high performance solution to M&S of confrontation operations

For future work, we will further explore the feasibility of quantifying the degree of panic of a crowd. We are also interested in the approaches to detecting groups in a crowd and identification of the leader of a group.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grant No. 60804036, 90820016), National Science Fund for Distinguished Young Scholars (61025019), the Hundred University Talent of Creative Research Excellence Programme (Hebei, China), and the Fundamental Research Funds for the Central Universities (CUGL100608, CUG, Wuhan, and the Programme of High-Resolution Earth Observing System (China)). D. Chen gratefully acknowledges support from the Birmingham-Warwick Science City Research Alliance.

References

1. Bian C, Chen D, Wang S 2010. Velocity field based modelling and simulation of crowd in confrontation operations. In: Proceedings of the 16th IEEE international conference on parallel and distributed systems, pp 646–651
2. Chen D, Theodoropoulos GK, Turner SJ, Cai W, Minson R, Zhang Y (2008) Large scale agent-based simulation on the grid. Future Gener Comput Syst 24(7):658–671
3. Chen D, Li D, Xiong M, Bao H, Li X (2010) GPGPU-aided ensemble empirical mode decomposition for electroencephalogram analysis during anaesthesia. IEEE Trans Inf Technol Biomed 14(6):1417–1427
4. Chen D, Turner SJ, Cai W, Theodoropoulos GK, Xiong M, Lees M (2010) Synchronization in federation community networks. J Parallel Distrib Comput 70(2):144–159
5. Chen D, Wang L, Ouyang G, Li X (2011) Massively Parallel Neural Signal Processing on a Many-core Platform. IEEE Comput Sci Eng. doi:10.1109/MCSE.2011.20
6. Chenney S (2004) Flow tiles. In: Proceedings of the 2004 ACM SIGGRAPH/Eurographics symposium on computer animation, pp 233–242
7. Helbing D, Farkas I, Vicsek T (2000) Simulating dynamical features of escape panic. Lett Nat 407:487–490
8. Herrmann-Pillath C (2010) Entropy, function and evolution: naturalizing peircian semiosis. Entropy 12(2):197–242
9. Huang L, Wong SC, Zhang M, Shu C-W, Lam WHK (2009) Revisiting Hughes' dynamic continuum model for pedestrian flow and the development of an efficient solution algorithm. Transp Res, Part B, Methodol 43(1):127–141
10. Hughes RL (2002) A continuum theory for the flow of pedestrians. Transp Res, Part B, Methodol 36(6):507–535
11. JCuda (2011) http://www.jcuda.org/

12. Liu C, Hu B 2009. Mutual information based on Renyi's entropy feature selection. In: Proceedings of the IEEE international conference on intelligent computing and intelligent systems 2009, pp 816–820
13. Schwarz G, Mosler HJ (2005) Investigating escalation processes in peace support operations: An agent-based model about collective aggression. In: Proceedings of the third annual conference of the European social simulation association
14. Testa B (2009) Dispersal (entropy) and recognition (information) as foundations of emergence and dissolvence. Entropy 11(4):993–1000
15. Wang L, Fu C (2010) Research advances in modern cyberinfrastructure. New Gener Comput 28(2):111–112
16. Wang L, Jie W (2009) Towards supporting multiple virtual private computing environments on computational grids. Adv Eng Softw 40(4):239–245
17. Wang L, von Laszewski G, Tao J Kunze M (2009) Grid virtualization engine: design, implementation and evaluation. IEEE Syst J 3(4):477–488
18. Wang L, von Laszewski G, Chen D, Tao J, Kunze M (2010) Provide virtual machine information for grid computing. IEEE Trans Syst Man Cybern 40(6):1362–1374
19. Wang L, von Laszewski G, Kunze M, Tao J (2010) Cloud computing: a perspective study. New Gener Comput 28(2):137–146
20. Wang L, von Laszewski G, Kunze M, Tao J, Dayal J (2010) Provide virtual distributed environments for grid computing on demand. Adv Eng Softw 41(2):213–219
21. Zhou SP, Chen D, Cai WT, Luo LB, Low YH, Tian F, Tay SH, Ong WS, Hamilton BD (2010) Crowd modeling and simulation technologies. ACM Trans Model Comput Simul 20(4). doi:10.1145/1842722.1842725