A Movement Data Analysis and Synthesis Tool for Museum Visitors' Behaviors

Kingkarn Sookhanaphibarn and Ruck Thawonmas

Intelligent Computer Entertainment Laboratory
Department of Human and Computer Intelligence
Ritsumeikan University
Kusatsu, Shiga 525-8577, Japan

Abstract. Achievement of museum guide systems, in physical and virtual worlds, providing the personalization and context awareness features requires the prior analysis and identification of visitors' behaviors. This paper analyzes and synthesizes visitors' behaviors in museums and art galleries by using our defined parameters. A visit time and a observation distance can be calculated by using the proposed functions. The proposed synthesis algorithm is developed and used in classification. Classifying visitor styles is simply implemented by using the average and variance of their stopover time at and distance to all exhibits as shown in this paper.

Keywords: Visitor behavior, Movement style, Art gallery, Digital Museum, Data synthesis and analysis, User classification.

1 Introduction

A huge number of emerging multimedia technologies used in museums offers new opportunities of various presentations, where visitors can absorb the information of exhibitions in order of their preferences. Visitors in an art gallery have their stereotypical movement, which is categorized in four styles as proposed by Veron and Levasseur [1]. Four visiting styles based on animals' behavior are ant, fish, grasshopper, and butterfly styles. The ant visitor spends quite a long time to observe all exhibits by walking closer to exhibits but avoids empty spaces. The fish visitors prefer to move and stop empty spaces but avoid areas near exhibits. The grasshopper visitors spend a long time to see selected exhibits but ignore the rest of exhibits. The butterfly visitors observe almost exhibits but spend varied times to observe each exhibit. Identifying their visiting styles can take advantage of a guide system in museums as mentioned in [2,3,4,5].

In this paper, we propose an analysis approach and a synthesis function of four visitor styles based on their definitions, which are summarized in [1,6]. This is based on our hypothesis that a visiting style can be simulated by using mathematic functions. To validate our hypothesis, we conduct a classification using synthesized visiting data obtained by using mathematic functions. The contributions of this work are (a) the first analysis of this kind on museum-visitor

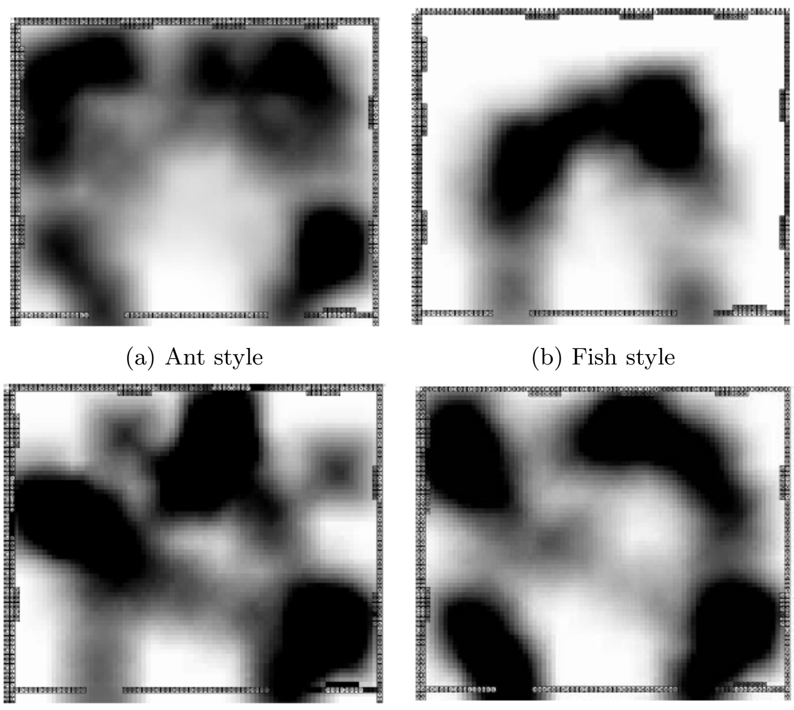
P. Muneesawang et al. (Eds.): PCM 2009, LNCS 5879, pp. 144–154, 2009.

visualization, (b) the novel synthesis method of visiting styles in art galleries and museums, (c) the classification approach in the form of their average and variance of stopover times, and (d) our implications to applications of these findings.

2 Visitors' Behaviors in Museums

Visitors' behaviors in artistic environments have received attention, since Veron and Levasseur [1] conducted their researches based on ethnographic studies in 1983. Their findings about four visiting styles have been cited in a mobile museum guide [2,3,4], a nomadic information system [7], a visualization tool [6], a museum audiovisual narration [8] . Bianchi et.al. [2], Gabrielli et.al. [3], Zancanaro et.al. [4], and Hatala et.al. [5] analyzed and developed a museum guide system via Personal Digital Assistants (PDA) technologies. Oppermann and Specht [7] proposed a prototype of a context-sensitive nomadic information system, whose the information can be adapted to user's knowledge and preferences.

Chittaro and Ieronutti [6] described four visiting styles based on results from their visualization tool, where black highlights areas more traveled, white identifies



(a) Ant style (b) Fish style

(c) Grasshopper style (d) Butterfly style

Fig. 1. Visualization of visitor styles excerpted from *A Visual Tool for Tracing Users' Behavior in Virtual Environments* by Chittaro et.al. in 2004

the areas less traveled and different shades of gray are used to identify intermediate previous situations, follows:

1. In visualization of an ant visitor, the center of most space on the map is colored in black, and there are no large differences in the time different exhibits have been seen as shown in Fig. 1 (a).
2. In visualization of a fish visitor, the areas near all exhibits on the map are colored in black as shown in Fig. 1 (b).
3. In visualization of a grasshopper visitor, areas near some exhibits are colored with highly variable shades of gray due to the fact that this visitor spends a variable time to observe different exhibits and ignores the rest of them as shown in Fig. 1 (c).
4. In visualization of a butterfly visitor, the areas near all exhibits are colored with different shades of gray and some are colored with black, but less regularly than those of the ant visitors, as shown in Fig. 1 (d).

Note that Fig. 1 shows the visualization of a visitor in the art gallery, where all exhibits are placed on the wall.

The aforementioned researches [1,3,6] described four visiting styles beneficial to a physical environment design in museums as follows:

1. Ant visitors need to be guided by a powerful rationale and this is the reason why they usually follow the path proposed by the museum curator.
2. Fish visitors prefer a holistic observation area, such as the center of the room.
3. Grasshopper visitors move directly to their selective exhibits.
4. Butterfly visitors refuse to follow the given path designed by other people and prefer their own route.

2.1 Analysis of Visitors' Behaviors in Museums

Considering the visualization scheme of visiting styles in [6], we can discard the temporal information in trajectories data of visitor movement in a museum. Given a visit map of $m \times n$ pixels, it is obvious that a dark colored area is a key role to identify the visiting styles because it illustrates a long stopover. Suppose that a museum area, $U = \{(i_1, i_2) | 1 \leq i_1 \leq m,\ 1 \leq i_2 \leq n\}$, is a set of pixels. The patterns of dark colored area can identify the visiting styles, and are related to the following parameters:

1. Visit time $v(\mathbf{i})$ is defined as the visitor's stopover time at pixel $\mathbf{i}$
2. Visit map is defined as a set of $v(\mathbf{i})$, where $\mathbf{i} \in U$
3. Observation distance $o(\mathbf{i})$ is defined as the visitor's fuzzy distance from the nearest exhibit
4. Observation map is defined as a set of $o(\mathbf{i})$, where $\mathbf{i} \in U$

Let $\mathbf{h}_k$ be the position of exhibit k and let $H = \{\mathbf{h}_k | 1 \leq k \leq M\}$ be a set of all positions of exhibits, where M is a total exhibit. Let $\mathbf{u}$ be the position of a visitor. The exhibit $\mathbf{h}_{k^*}$ nearest to visitor $\mathbf{u}$ is calculated by the following

$$k^* = \text{argmin}_k \|\mathbf{h}_k - \mathbf{u}\| \quad \textit{where} \quad 1 \leq k \leq M \tag{1}$$

where $\|\mathbf{h}_k - \mathbf{u}\|$ is the Euclidean distance between exhibit $\mathbf{h}_k$ and visitor $\mathbf{u}$.

To compute the observation distance, the observation distance function $\Psi(\mathbf{h}_k, \mathbf{u})$ is defined as an exponential function of the distance from exhibit $\mathbf{h}_k$ to visitor $\mathbf{u}$ as follows:

$$\Psi(\mathbf{h}_k, \mathbf{u}) = e^{-\rho\|\mathbf{h}_k - \mathbf{u}\|^2} \tag{2}$$

ρ is a weighted constant. In the same manner of an Gaussian-based activation function, this weighted constant is obtained from the variance of the Euclidean distance between exhibit $\mathbf{h}_k$ and pixel $\mathbf{i}$, for all $\mathbf{h}_k$ and $\mathbf{i}$, as shown in Eq. 3.

$$\mu = mean\left\{\bigcup_{\mathbf{h}_k \in H \,\wedge\, \mathbf{i} \in U} \|\mathbf{h}_k - \mathbf{i}\|\right\} \tag{3a}$$

$$\sigma^2 = mean\left\{\bigcup_{\mathbf{h}_k \in H \,\wedge\, \mathbf{i} \in U} (\|\mathbf{h}_k - \mathbf{i}\| - \mu)^2\right\} \tag{3b}$$

$$\rho = 2\sigma^2 \tag{3c}$$

Fig. 2 shows a graph plotting of $\Psi(\mathbf{h}_i, \mathbf{u})$ and $\|\mathbf{h}_k - \mathbf{v}\|$ where σ^2 is 0.07. The observation distance is calculated as follows:

$$o(\mathbf{i}) = \Psi(\mathbf{h}_k, \mathbf{i}) \quad where \quad \mathbf{i} \in S(\mathbf{h}_k) \tag{4}$$

The visitor's attendance to exhibit $\mathbf{h}_k$ is computed by using the visit time and the observation distance belonging to $S(\mathbf{h}_k)$, as shown below:

$$A(\mathbf{h}_k) = mean\left\{\bigcup_{\mathbf{i} \in S(\mathbf{h}_k)} \{o(\mathbf{i})\, v(\mathbf{i})\}\right\} \tag{5}$$

where $S(\mathbf{h}_k)$ is a set of pixels belonging to $\mathbf{h}_k$, as defined by Eq.(1), and $o(\mathbf{i}) \geq \delta$. δ is a threshold in the range from 0 to 1, which its implication is a observation distance limit.

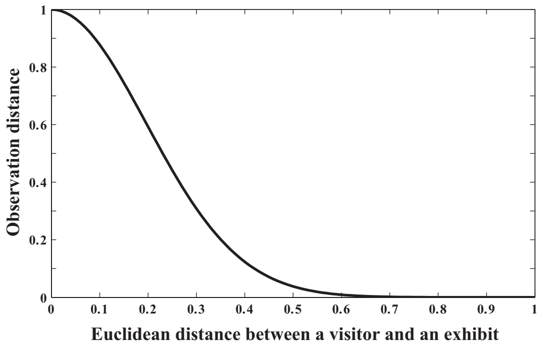


Fig. 2. Observation distance function of Eq.(2)

3 Synthesis of Visitors' Behaviors in Museums

To synthesize the visit map, there are two input data as follows: a plan of museum, which illustrates the location of all exhibits, and his/her visitor type. We propose a synthesis approach for four visiting styles through the following steps:

1. Selective exhibit randomness
2. Preference weight randomness
3. Visit time distribution
4. Noise synthesis
5. Smoothing filter

The difference between the grasshopper and the butterfly styles is the number of stopovers at exhibits, i.e., the former stops fewer than the latter does. The selective exhibit randomness step is used in these two visiting styles. The preference weight randomness step generates varying stopover times at exhibits. Both ant and fish styles show no significant difference among their preferences of exhibits. Second, the visit time distribution functions are different among visiting styles. Third, noise synthesis is conducted using Gaussian Probability Distribution Function (PDF). The implication of noise is randomly short stops in the exhibition area. Then, the smooth filtering such as a low pass filter is applied in order to discard the sharp detail and remove the noise.

The following notations are used in four synthesis algorithm; the preference weight of the exhibit $\mathbf{h}_i$ is denoted by $w(\mathbf{h}_i)$, and ρ is a weighted constant. Without loss of generality, an art gallery shown in Fig. 3 is used as the museum map in our synthesis, which there are 12 exhibits hanging in three sides of the room.

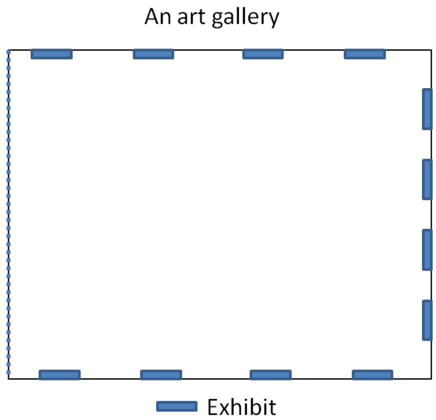


Fig. 3. The plan of an art gallery which all exhibits are hung on the wall

3.1 Ant Visiting Style

The ant visitor always stands closest to every exhibit because he/she behaves like a curious visitor pays attention to what he/she is looking at. He/she skips none of exhibits and takes a long stop time at every exhibit. That is the reason why a distribution function of the visit time is defined as the same shape of the observation distance function as shown in Fig. 2. Our synthesis algorithm of the *Ant* visit map, of which darkest colored pixels equal to 1, is shown as follows:

1. Preference weight randomness is set by using the Gaussian PDF with the mean closer to 1 and smallest variance.
2. Visit time $v(\mathbf{i})$ at pixel $\mathbf{i}$ is defined by using the following equation:

$$v(\mathbf{i}) = w(\mathbf{h}_k)e^{-\rho\|\mathbf{h}_k-\mathbf{i}\|^2} \text{ where } \mathbf{i} \in S(\mathbf{k}) \tag{6}$$

3. Smoothing filter is applied to the visit map obtained from Step 2.

The selective exhibit randomness and the noise synthesis steps are not applied because the ant visitor stops at all exhibits and moves following the exhibit path proposed by a museum curator. An example of the ant visit map derived by using the above algorithm is shown in Fig. 4.

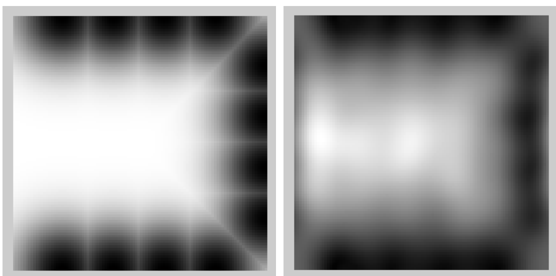
3.2 Fish Visiting Style

The fish visitor always spends much of his/her time in the empty space. On the other hand, he/she does not stand close to any exhibit. That is the reason why a distribution function of the visit time is defined as the inverse shape of the observation distance function. Our synthesis algorithm of the *Fish* visit map is shown as follows:

1. Preference weight randomness is set by using the Gaussian PDF with the mean closer to 1 and smallest variance.
2. Visit time $v(\mathbf{i})$ at pixel $\mathbf{i}$ is defined by using the following equation:

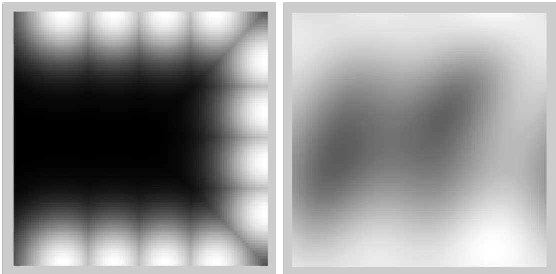
$$v(\mathbf{i}) = w(\mathbf{h}_k)(1 - e^{-\rho\|\mathbf{h}_k-\mathbf{i}\|^2}) \text{ where } \mathbf{i} \in S(\mathbf{k}) \tag{7}$$

3. Smoothing filter is applied to the visit map obtained from Step 2.



(a) PDF Generation (b) Smoothing filtering

Fig. 4. Visualization of an *ant* visitor



(a) PDF Generation (b) Smoothing filtering

Fig. 5. Visualization of a *fish* visitor

The selective exhibit randomness and the noise synthesis are not applied because the fish visitor stops at none of exhibits and moves only to the center of the room. An example of the fish visit map derived by using the above algorithm is shown in Fig. 5.

3.3 Grasshopper Visiting Style

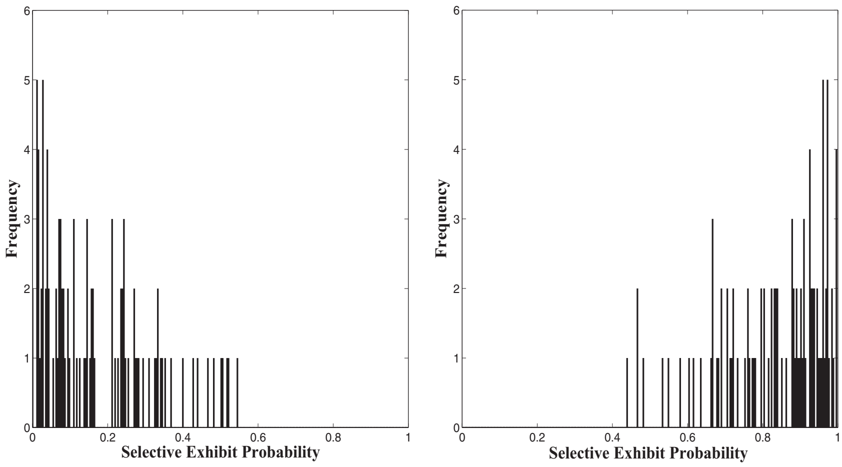
The grasshopper visitor prefers to stand close to a few exhibits in which he/she has an interest and ignores the remaining part. Unlike the ant style, a grasshopper distribution function of the visit time has an addition parameter, i.e., an selective exhibit f_k. The f_k of exhibit k by which the grasshopper visitor is attracted is assigned to 1. If f_k equals to 1, the distribution function of the visit time is defined as the same function of the ant style. Otherwise, the visit time is set to zero. Our synthesis algorithm of the *Grasshopper* visit map is shown as follows:

1. Selective exhibit randomness is decided by using Beta PDF with $\gamma = 1$ and $\beta = 5$. Given sample $N = 100$, the Beta PDF is shown in Fig. 6(a). The output of the PDF is the number of selected exhibits, as denoted by L. The selective exhibit is defined as a binary vector where f_i is an element of the binary vector and $\sum_i f_i$ must be L.
2. Preference weight randomness is set by using the Gaussian PDF with the mean closer to 1 and small variance.
3. Visit time $v(\mathbf{i})$ at pixel $\mathbf{i}$ is defined by using the following equation:

$$v(\mathbf{i}) = \begin{cases} w(\mathbf{h}_k)e^{-\rho\|\mathbf{h}_k - \mathbf{i}\|^2} & \text{if } \mathbf{i} \in S(\mathbf{k}) \text{ and } f_k = 1 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

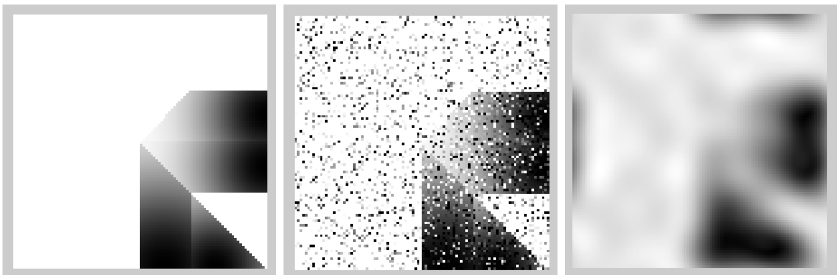
4. Noise synthesis is performed by using Gaussian PDF.
5. Smoothing filter is applied to the visit map obtained from Step. 4.

An example of the grasshopper visit map derived by using the above algorithm is shown in Fig. 7.



(a) $\gamma = 1$ and $\beta = 5$ (b) $\gamma = 5$ and $\beta = 1$

Fig. 6. Histogram of the Beta PDF by given sample $N = 100$ where the PDF functions of selective exhibit randomness of (a) grasshopper and (b) butterfly



(a) PDF (b) After adding noise (c) After smoothing filter

Fig. 7. Visualization of a *grasshopper* visitor

3.4 Butterfly Visiting Style

The butterfly visitor behaves like the ant visitor but spends much varied stop time at each exhibit. He/she sometimes ignores a few exhibits. On the other hand, the butterfly visitor has the same distribution function of the visit time but the different randomness function of the selective exhibit parameter, f_k. Our synthesis algorithm of the *Butterfly* visit map is shown as follows:

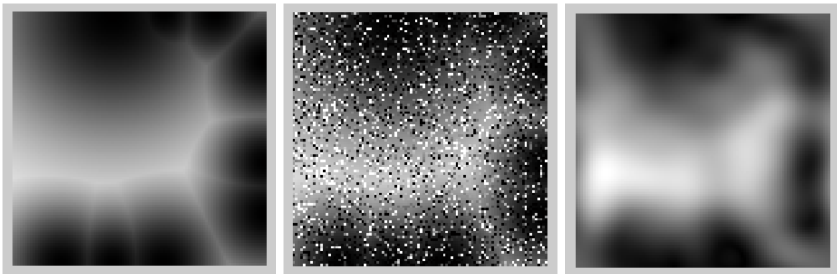
1. Selective exhibit randomness is decided by using Beta PDF with $\gamma = 5$ and $\beta = 1$. Given sample $N = 100$, the Beta PDF is shown in Fig. 6(b). The output of the PDF is the number of selected exhibits, as denoted by L. The selective exhibit is defined as a binary vector where f_i is an element of the binary vector and $\sum_i f_i$ must be L.

2. Preference weight randomness is set by using the Gaussian PDF with the mean closer to 1 and variance higher than the ant's.
3. Visit time $v(\mathbf{i})$ at pixel $\mathbf{i}$ is defined by using the following equation:

$$v(\mathbf{i}) = \begin{cases} w(\mathbf{h}_k)e^{-\rho\|\mathbf{h}_k-\mathbf{i}\|^2} & \text{if } \mathbf{i} \in S(\mathbf{k}) \text{ and } f_k = 1 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

4. Noise synthesis is performed by using Gaussian PDF.
5. Smoothing filter is applied to the visit map obtained from Step. 4.

An example of the butterfly visit map derived by using the above algorithm is shown in Fig. 8.



(a) PDF (b) After adding noise (c) After smoothing filter

Fig. 8. Visualization of a *butterfly* visitor

4 Classification

Here, we propose a simple classification approach derived by an average and variance of $A(\mathbf{h}_i)$ for all i. The experiments are designed into two phases as follows: the data synthesis and the classification. We synthesized is a total of 100 visitors, whose type is randomly defined by the uniform distribution. The number of ant, fish, grasshopper, and butterfly visitors are 29, 24, 26, and 21, respectively. Given a museum size of 100×100 pixels, the position of all exhibits is shown by the following two-dimensional vectors:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 20 & 40 & 60 & 80 & 100 & 100 & 100 & 100 \\ 20 & 40 & 60 & 80 & 100 & 100 & 100 & 100 & 20 & 40 & 60 & 80 \end{bmatrix}^T$$

The visit maps of ant, fish, grasshopper, and butterfly visitors were generated by using the algorithms in Sec. 3. The observation map of the museum was also synthesized by using the observation distance function in Eq.(2). The visitor types can be clustered based on their averages and variances of the visitor's attendance to all exhibits, which are calculated by using Eq.(5). The clusters of four visitor types as shown in Fig. 9 indicate that the clusters of ant and fish visitor types are isolated, which their probabilities found in [2] are 30% and 20%,

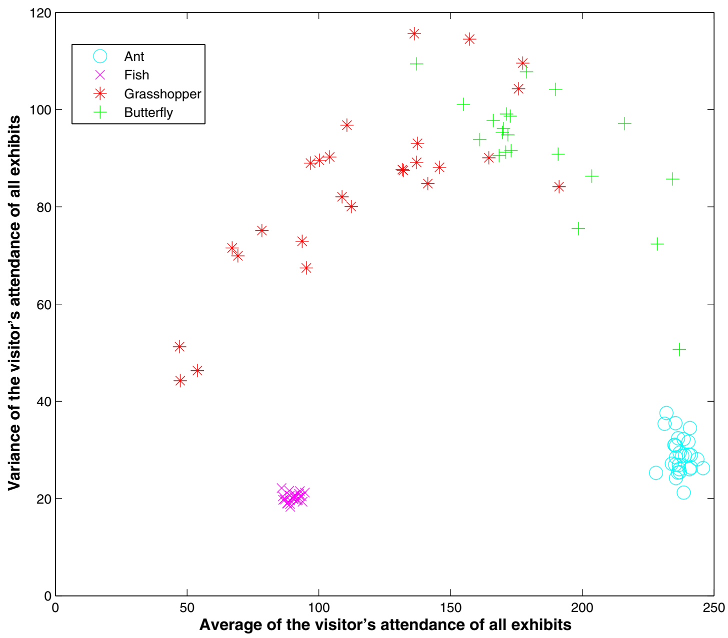


Fig. 9. A scatter graph of 100 visitors between their averages and variances of $A(\mathbf{h}_i)$ for all i

respectively. The selective exhibit randomness is a cause of an overlapping area between grasshopper and butterfly clusters with the result that the Beta PDFs of both visitor types are a common area. However, the probability of grasshopper visitors found in the physical world is only 10% as reported in [2]. Other intelligent computing techniques, such as fuzzy systems and neural networks, can be used in classification among the grasshopper and butterfly visiting types.

5 Conclusions and Discussions

The findings in this paper are parameters related to visitors' styles and their functions, which are used in classification. The visit time and observation distance are described in the paper. In classification, our method is simple to understand and implement. Four visiting styles considered -in this paper are ant, fish, grasshopper, and butterfly. Our synthetic data of those styles were derived from visitors' preference and interest in the museums and art galleries. Besides, there are other influences over visitors such as their company. The visitors' behavior can be substantially changed when they accompany their relatives or friends on the museum trip. Therefore, the effects of visitor interaction on the visiting styles will be investigated in the future work.

Our synthesis algorithm will aim to visualize an available and crowed space in art galleries and museums, where a particular proportion of visiting styles and a time duration are provided. This tool will be useful as shown in the following fields: (a) Computer Aided Design (CAD) for a physical environment in art galleries as addressed in [3,6]. (b) Content Management System (CMS) for digital museums in a virtual world as addressed in [9,10], such as Second Life emerged as a massively multiuser online world supporting about 60,000 simultaneous login users.

Acknowledgment

This work is supported in part by Global COE (Center of Excellence) Program "Digital Humanities Center for Japanese Arts and Cultures" at Ritsumeikan University.

References

1. Veron, E., Levasseur, M.: Ethnographie de l'Exposition. Bibliothque publique d'Information, Centre Georges Pompidou, Paris (1983)
2. Bianchi, A., Zancanaro, M.: Tracking users' movements in an artistic physical space. In: The i^3 Annual Conference, Siena, Italy, October 20-22, pp. 103–106 (1999)
3. Gabrielli, F., Marti, P., Petroni, L.: The environment as interface. In: The i^3 Annual Conference, Siena, Italy, October 20-22, pp. 44–47 (1999)
4. Zancanaro, M., Kuflik, T., Boger, Z., Goren-Bar, D., Goldwasser, D.: Analyzing museum visitors' behavior patterns. In: Conati, C., McCoy, K., Paliouras, G. (eds.) UM 2007. LNCS (LNAI), vol. 4511, pp. 238–246. Springer, Heidelberg (2007)
5. Hatala, M., Wakkary, R.: Ontology-based user modeling in an augmented audio reality system for museums. User Modeling and User-Adapted Interaction (3-4), 339–380 (2005)
6. Chittaro, L., Ieronutti, L.: A visual tool for tracing users' behavior in virtual environments. In: AVI 2004: Proceedings of the working conference on Advanced visual interfaces, pp. 40–47. ACM, New York (2004)
7. Oppermann, R., Specht, M.: Context-sensitive nomadic exhibition guide. In: Proc. Ubicomp 2000, Berlin, pp. 127–142. Springer, Heidelberg (2000)
8. Sparacino, F.: The museum wearable: real-time sensor-driven understanding of visitors' interests for personalized visually-augmented museum experiences. In: Proceedings of Museums and the Web (MW 2002), pp. 17–20 (2002)
9. Sookhanaphibarn, K., Thawonmas, R.: A framework for design and evaluation of digital museums in second life as learning institutions. In: Proceedings of International Conference on Information Hiding and Multimedia Signal Processing (IIH-MSP). IEEE, Los Alamitos (2009)
10. Sookhanaphibarn, K., Thawonmas, R.: A content management system for user-driven museums in second life. In: Proceedings of International Conference on Cyberworld. IEEE, Los Alamitos (2009)