Mood-Learning Public Display: Adapting Content Design Evolutionarily through Viewers’ Involuntary Gestures and Movements

Ken NAGAO†(a), Nonmember and Issei FUJISHIRO†(b), Member

SUMMARY Due to the recent development of underlying hardware technology and improvement in installing environments, public display has been becoming more common and attracting more attention as a new type of signage. Any signage is required to make its content more attractive to its viewers by evaluating the current attractiveness on the fly, in order to deliver the message from the sender more effectively. However, most previous methods for public display require time to reflect the viewers’ evaluations. In this paper, we present a novel system, called Mood-Learning Public Display, which automatically adapts its content design. This system utilizes viewers’ involuntary behaviors as a sign of evaluation to make the content design more adapted to local viewers’ tastes evolutionarily on site. The system removes the current gap between viewers’ expectations and the content actually displayed on the display, and makes efficient mutual transmission of information between the cyberworld and the reality.

Key words: public display, image processing, human behavior recognition, genetic algorithm

1. Introduction

In order to deliver the original message effectively, any signage is required to provide “attractive” content to its potential viewers. To that end, it is important to survey viewers’ evaluations towards the signage and to modify the appearance of the content accordingly. In general, such survey and modification are carried out manually because it is necessary to allow for the subjective matters as well as aesthetic issues. However, in actual signage, these manual modifications take time, and bring on the gap between viewers’ expectations and the content actually shown on the signage.

Recently, public display has attracted more attention as a new type of signage. One of the most important characteristics of public display is that its content can be replaced on site, which enables the instant adaptation of the content by utilizing interactions with viewers. A typical example is a public display with touchscreen. However, even in the methods utilizing such interactive devices, there remains an efficiency problem that these systems need to wait for viewers’ voluntary interactions. Some attempts address the problem that most viewers tend to ignore the public display, which is referred to as “display blindness”.

On the other hand, viewers also involuntarily behave in various ways in front of the public display system. For example, if the content is attractive, a viewer may stop to read the content, and he may change his head angle for reading each part of the content. If the part was convincing he would nod or if it was unacceptable he would shrug his shoulders. In these ways, through many contexts we can understand viewers’ mood towards the contents, while there have been few studies on utilizing these behaviors for public display systems so far.

From this aspect, we herein propose Mood-Learning Public Display as a public display system which estimates viewers’ mood towards the content design from their involuntary behaviors and uses the mood for automatic, instant adaptation of the content design to local viewers. The adaptation method for content designs utilizes an idea of genetic algorithm (GA). The system can gather evaluations from a number of viewers without any need to wait for their voluntary interactions. Additionally, those behaviors come directly from their actual feeling towards the content designs, and thus it is not necessary to check up the validity of the evaluation.

In our previous papers [1], [2] and poster [3], we focused primarily on viewer’s specific gestures, such as folding arms and nodding, while the viewer’s movements in front of a public display reflect the viewer’s mood towards the content directly. Therefore, our new system analyzes viewers’ mood towards the content design from various contexts including viewers’ movements as well. In order to empirically evaluate the utility of the method, we have applied a prototype of Mood Learning Public Display to the display of conference poster. The proposed system in actual use can be seen in Fig. 1, and Fig. 2 shows an example of adaptive conference poster in Japanese. Figure 4 illustrates the processing flow of the proposed method and how the viewers’ behaviors are recognized (see details in Sect. 3 and Sect. 4).
2. Related Work

2.1 Classification of Previous Systems

As we mentioned, it is important to survey viewers’ evaluations towards the signage and to modify the appearance of the content accordingly. So far, direct questions to viewers, such as interviews or questionnaires, have been commonly used for evaluating the content of public display [4]. One major drawback in such direct questions is that much human labor is required for gathering and analyzing the sufficient data of evaluation. In addition to this, it is necessary to check up the validity of the answers because they do not always tell their actual feeling in such enforced situations.

On the other hand, recent developments of hardware technology have been changing the style of content evaluation and modification. As previously discussed, public display attracted much attention as a new type of signage, and one of its most important characteristics is that the content can be replaced on site. Here, we categorize public display systems according to “what input information the system mainly uses for changing its content”. Figure 3 shows a classification tree of public display systems.

As can be seen from Fig. 3, public display systems can be firstly categorized by whether they utilize information mainly through viewers or environment. The secondary aspect is whether the information is dynamic or static. Based on the classification, we will survey previous systems and methods for public display, with special emphasis on “Information through viewers”, which are summarized in Table 1.

2.2 Utilization of Viewers’ Appearance

There are many studies on obtaining information through the viewer’s appearance, especially through face detection [5], [6], and many products for public display are commercially-available. These R&Ds enable the public display system to get viewers’ demographic segmentations, such as sex and age, while the information does not include viewers’ attitudes towards public display, and thus the system cannot estimate what viewers really want to see.

2.3 Utilization of Viewers’ Behaviors

Some public display systems require voluntary behaviors of viewers so that they can understand what viewers want to see, which are especially useful when specific input from viewers are needed. Recent development of interactive devices popularizes such public display systems. One of the most general interactive devices is an external input device such as keyboard and/or a mouse. Additionally, touchscreen made public displays released from the use of such external input devices, allowing wide workspace and intuitive interfaces. However, viewers may not find out that the system is interactive when there is no external input device. Recently, some public displays start to utilize non-contact devices for interactions, such as mobile phones, and there are some systems which do not even require such devices for interactions. Such systems take advantage of motion sensor devices such as Microsoft Kinect™, whose reliability is verified through previous attempts of R&D [16]. These systems allow the viewers to change their content based on their gestures, such as hand swipe motion.

The above systems are utilizing viewers’ voluntary actions, while viewers’ involuntary behaviors are also of great significance. Vogel et al. [9] considered the viewers’ behaviors from the aspect of their relative positions and the aspect of their voluntary behaviors of touching screen, for changing levels of detail of the content. Schiavo et al. [10] proposed

![Fig. 2](image-url) An example of adaptive conference poster displayed by the proposed system.

![Fig. 3](image-url) Classification tree of public display systems from the aspect of main input for changing their content.

| Table 1 Classification of previous public display studies. |
|-----------------------------------------------------------|
| **Public display**                                        |
| Dynamic behaviors                                       |
| Voluntary behaviors | Involuntary behaviors | Static appearance |
| Body gesture [7]–[9] w/ mobile phone [11]–[13] w/ touchscreen [9], [11], [12] w/ keyboard [14] w/ mouse [15] |
| Approaches [7], [9], [10] Passing-by [7], [8] Orientations [9], [10] |
| Face detections [11] |
the method to measure how much attention the viewer pays to the public display from his behaviors such as position and orientation. In addition, Muller et al. [8] made the system which projects viewers’ passing-by motions onto display, in order to attract viewers. The problem that viewers do not recognize the public display is referred to as the problem of “display blindness” and widely known. As to methods to attract attention from viewers, Khan et al. [17] utilized light to attract attention from viewers. Most of the previous systems utilizing viewers’ involuntary behaviors are focusing only on fundamental behaviors, such as positions and body orientations, and there is few prior works which utilize more various behaviors to provide an integrated framework for evaluating viewers’ involuntary behaviors.

2.4 Modification of Content Design

As to the modification of the content design, most content modification of the previous systems is only displaying pre-defined content designs, and does not focus on seeking other better content designs. The question on the attractive content design has been a difficult question because it involves many subjective issues ranging from psychology to other disciplines. There have been some research attempts to tackle this problem. For example, Singh and Bhattacharya proposed an algorithm to improve the aesthetics of Web interface utilizing a GA [18], using more than ten geometrical features of page layout as the evaluations. However, the approach does not take users’ evaluation to the Web interface into primary account, and they concluded that it is difficult to define the cases in which the approach works well. Moreover, in the case of public display, it should be considered that local tastes towards the content may be different depending on the place where it is installed.

2.5 Standpoint of Our Approach

Considering the pros and cons of these methods, we have proposed an idea of making public display more attractive by utilizing viewers’ specific involuntary behaviors with the idea of a GA [1]–[3], which can be categorized as the one to use “involuntary behaviors” in Fig. 3. To the best of our knowledge, these were only attempts to utilize viewers’ specific involuntary behaviors for automatically making public display more attractive. However, the method only considers viewers’ involuntary gestures, while their involuntary positional movements also have big implications. For example, if the viewer is viewing far from the public display, this means that he is having a skim through the content. On the other hand, if he is taking a look at the content near the public display, he may be taking his time to read the content. Also, approaching the content means that he finds an interesting part of the content. Our new public display system also detects these as important signs of viewers’ degree of interest.

3. Approach

As shown in Fig. 4, the main processing flow in our system is composed of two steps: (a) evaluation step and (b) modification step. Hereafter, we will refer to a person who uses this system for delivering his messages to viewers simply as “sender”. Repeating these steps, the system evolutionarily produces more attractive content designs, and automatically reaches an ideal content design. Here, “ideal” means that it fits local tastes, with an assumption that different local viewers are likely to have similar moods towards similar designs. The system terminates the evolution when many similar designs appear or when a large number of viewers attend and
they become less likely to behave in a negative way. Our experiments show that around seventh generation is mostly enough, but the detailed conditions of the termination can be determined by the sender.

Figure 2 illustrates one example of evolved content designs. In the evolved content design, each of the figures in the poster is independently adjusted in terms of size, and each section is repositioned based on its relationships with the others. Also, some sections are modified for its attractiveness in the evolved content design.

3.1 Defining a Content

Firstly, the sender is required to define what he wants to show to the public, dividing the content into smaller sections. For example, an academic conference poster can generally be divided into “title”; “authors”; “section titles”; “content of sections”; and “reference figures”. Hereafter, we will simply refer to the overall design of the content as content design, and small content sections defined in this step as sections.

Each section is required to have two kinds of properties: content properties and graphical properties. Content properties are what the sender wants to tell to the public in the corresponding section. For example, in the section of “authors”, content properties are specified as the names of the authors. On the other hand, graphical properties indicate decoration of the section, such as font; font color; frame color; and size of the section. Based on these properties, the system automatically estimates the relationships between these gestures and feelings which are deeply related to the viewers’ emotions. Finding the relationships between these gestures and feelings has been studied extensively [19]. Our system can learn the meaning of the gestures referring to such a kind of database of our own that maintains the relationships. Here, we categorized fifteen varieties of behaviors, including the gestures we have mentioned.

The movements of viewers are represented with “current positions” and “moving directions”. Utilizing the current position of the viewer, the system estimates whether the viewer is just having a skim through the content or concentrating on reading one part of the content. The moving direction of the viewer tells the system the viewer’s transition between the two reading styles, and indicates the viewer’s degree of interest towards the content.

When recognizing these behaviors, the system needs to take a difference in culture into account. Meanings of viewers’ behaviors may be different depending on cultures. Considering that previous research on the detailed meaning of gestures has limited capability because human affection is a subjective matter as mentioned before, we herein focus only on one culture, i.e., Japanese, and then classify the meanings just according to whether the gesture has positive or negative meaning.

3.2 Evaluation Step

3.2.1 Displaying Each Content Design

Firstly, the system starts to display one content design. Then, the content design will switch to another content design of the same generation when a fixed time duration runs over and at that point no viewer is looking at it. In this way, the system gathers evaluations for each content design of the same generation.

Here, in order to attract viewers’ attention and decrease the “display blindness”, the system illuminates the display brighter when passers-by are near the public display. This helps make the viewers realize the existence of the display, and can be regarded as a good trigger to initiate the following process to make the public display more attractive.

3.2.2 Recognizing Viewers’ Involuntary Behaviors

For a finite period of time, it is estimated that a number of viewers watch the public display one after another. Here, the system recognizes various behaviors of the viewers individually for the evaluation. The system secretly captures these behaviors through utilizing a capture device Microsoft Kinect™, which can capture behaviors of two viewers simultaneously. As we mentioned, our system estimates viewers’ mood from two aspects: “gestures” and “movements”, which can be captured through the capture device.

Here, we define two types of gestures: gestures indicating viewers’ attention point (attention pointer) and gestures indicating viewers’ feeling (feeling indicator). Attention pointers are gestures such as changing head angle and pointing gesture, which can be used for estimating viewers’ gaze points. On the other hand, feeling indicators are gestures such as folding arms; nodding; and shrugging shoulders, which are deeply related to the viewers’ emotions. Finding the relationships between these gestures and feelings has been studied extensively [19]. Our system can learn the meaning of the gestures referring to such a kind of database of our own that maintains the relationships. Here, we categorized fifteen varieties of behaviors, including the gestures we have mentioned.

The movements of viewers are represented with “current positions” and “moving directions”. Utilizing the current position of the viewer, the system estimates whether the viewer is just having a skim through the content or concentrating on reading one part of the content. The moving direction of the viewer tells the system the viewer’s transition between the two reading styles, and indicates the viewer’s degree of interest towards the content.

When recognizing these behaviors, the system needs to take a difference in culture into account. Meanings of viewers’ behaviors may be different depending on cultures. Considering that previous research on the detailed meaning of gestures has limited capability because human affection is a subjective matter as mentioned before, we herein focus only on one culture, i.e., Japanese, and then classify the meanings just according to whether the gesture has positive or negative meaning.

3.2.3 Analyzing Evaluations of Each Section

Based on these gathered data, the system surveys and analyzes the gathered evaluations of each section in two aspects: how attracting and unconvincing the section is.

The former one is defined as the total of weighted-time duration when viewers are looking at the section, no matter how they behave. We will refer to this evaluation as attracting evaluation of the section. The latter one is defined as the total of weighted-time duration when the viewers are looking at the section with negative feeling. We will refer to this evaluation as unconvincing evaluation of the section. Note that the system considers every gesture the viewer behaves and that the negative gesture is only considered for the time the gesture is behaved. These values are accumulated based on the individual viewers’ gaze points, and distributed to each section finally.

3.2.4 Analyzing Evaluations of Overall Structure

The system automatically estimates the relationships be-
tween the sections of the content, and does not require the sender to define it. The reason is that there may be a more ideal structure rather than the one the sender originally thought the content has. For example, in the case of academic conference posters, if there were many professional people in the field of the content near the public display, they would find the flowchart image as the reference image of the section “system overview”, while if most attendees were not conversant with the field so much, they would find the execution example images as the reference images of the section. An appropriate structure depends on a situation where public display is set, and the system needs to have an allowance for this. Because a viewer changes his gaze point based on the roles and meanings of the sections, the system surveys the sectional structure utilizing the trajectory of a viewer’s gaze point.

3.3 Modification Step

After all content designs are evaluated, they are sorted in an evaluation order based on the weighted-time duration each content design was looked at without unconvincing feelings. Then, using the above evaluations as a guide, the system generates the next generation of content designs in two ways: mutation and crossover, also inheriting the current top-ranked content designs. Here, because the weight of information each section has can be different, our system makes modifications to sections with heavy information such as the title smaller than other sections with light information, keeping the information inertia in mind. These amounts of information can be pre-defined by the sender. Figure 5 shows examples of the breeding.

3.3.1 Mutation

In this step, the system breeds some new individuals by changing the property values of each section in the highest-ranked content design. Since a section which got a high value of attracting evaluation is a section which the viewers found worth looking at, emphasizing the section is important for making the content design more attractive. Therefore, the system changes its graphical properties for the attractiveness. Note here that the system just randomly changes its graphical properties, because the unattractive individuals will be eliminated later and the only attractive individuals will naturally survive.

A section which got a high value of unconvincing evaluation is a section whose content property is thought to be unconvincing by the viewers, and thus the content properties of the section should be changed for making the content easier to understand, especially in the case the section got a high value of attracting evaluation. In the same way as before, the system only needs to randomly select the value from the pre-defined values of the content properties, because unconvincing individuals will be exterminated later.

Additionally, the system modifies the layout of content design based on the analysis about the relationships of sections. The system puts up a layout where mutually-related sections are set closer.

In Fig. 5, “system overview” (the section at the lower left) is emphasized because it gathered much attention from the viewers. On the other hand, the sentences of “approach” (the section at the lower right) are replaced with other pre-defined sentence patterns because the previous sentences were thought to be unconvincing by the viewers. Furthermore, the sections which were thought to have the relationships were set closer.

3.3.2 Crossover

In addition to those by mutation, the system breeds new individuals by crossover. From highly-ranked content designs, the system selects two content designs for crossover, and breeds new content designs by combining property values of each section in the two content designs. In Fig. 5, the system breeds a new content design whose properties come from the two previous content designs.

4. Algorithm

In this section, each of the steps outlined in Sect. 3 is detailed by describing its algorithms.

4.1 Defining a Content

When the total number of section is \( n \), each section is given one section number from 1 to \( n \) respectively. Here, the sender defines what values the content properties can take, because there exist various ways to express the message. On the other hand, graphical property values are given to each by the system automatically.

Giving random values to the graphical properties and randomly selecting the pre-defined values for the content properties of each section, the system creates a fixed number of initial content designs. Layout of each content design is also randomly set, while there is a constraint which requires the section “title” and “authors” to be set at the top of a poster so that each content design satisfies minimal requirements for a layout of an academic conference poster. These
generated content designs constitute the first generation of content designs. The sender can define the total number \( N \) of the content designs of the same generation. Here, the system gives each content design one content design number from 1 to \( N \), respectively.

4.2 Evaluation Step

For capturing the individual viewer’s behaviors, our system uses Microsoft Kinect™ and Kinect for Windows SDK 1.7, which allow the system to capture various behaviors such as head pose, facial expression, and gestures of multiple viewers involuntarily (See Fig. 4).

4.2.1 Analyzing Evaluations of Each Section

Attracting/unconvincing evaluations are approximated by Gaussian kernels separately. Here, the viewer’s gaze points in the content design, which can be estimated from the attention pointer, are regarded as the center of Gaussian distribution. Then, the evaluation value and radius of Gaussian distribution are determined by the degree of interest. When the viewer is getting closer, the radius is set small and the value becomes small. Figure 6 illustrates how viewers’ behaviors determines whether the evaluation is unconvincing, and the degree of interest affects the radius and value of the distribution. There is the other heightmap which indicates the attracting evaluation.

Example of gathering evaluation. The system uses Gaussian distributions and accumulates each evaluation by viewers until displaying the content design finishes. The viewer’s behavior determines whether the evaluation is unconvincing, and the degree of interest affects the radius and value of the distribution. There is the other heightmap which indicates the attracting evaluation.

4.2.2 Analyzing Evaluations of Overall Structure

In addition, the system recognizes the trajectories of viewers’ attention points and estimates the relationships between sections. Let \( F_j^v(x = 1 \cdots n) \) be a value which indicates how long a viewer has been focusing on the \( x \)-th section, where \( v \) indicates the identifier of the viewer given by the system. When the viewer \( v \) is looking at the \( f \)-th section, \( F_j^v \) increases according to the looking time, and when the section is not looked at, it decreases slowly until it comes to zero as time goes. In this way, the system memorizes which sections the viewer was looking at before.

On the other hand, each section has values which indicate the relationships with the other sections. Let \( R_j^k \) \((x = 1 \cdots n)\) be the values which the \( k \)-th section of the \( j \)-th content design has. When some viewer is looking at the \( f \)-th section when the \( j \)-th content design is displayed, \( R_j^f \) increase correspondingly based on \( F_1, F_2, \cdots, F_n \) the viewer has, and accumulate the values by every viewer. From these values, the system estimates the degrees of the relationships among sections. For example, when displaying the number \( j \) content design is finished, if \( \max[R_j^f, \cdots, R_j^m] \) is \( R_j^f \), this means that the \( f \)-th section has the biggest relationship with the \( m \)-th section in the \( j \)-th content design.

4.3 Modification Step

Here, let \( p_j^k \) be the probability that modification of graphical property happens at the \( k \)-th section of the \( j \)-th content design. This corresponds to the calculated attracting evaluation value of each section:

\[
p_j^k = \frac{E_j^k}{\max\{E_j^1, E_j^2, \cdots, E_j^n\}} \times \text{aggressiveness},
\]

where aggressiveness means how aggressive modifications are applied. This value can be defined by the sender. On the other hand, considering both attracting evaluation and unconvincing evaluation, a probability that the modification of content property is applied to the section (let \( p_j^{\prime k} \) be the probability in same way as \( p_j^k \)) can be formulated as follows:

\[
p_j^{\prime k} = \frac{U_j^k p_j^k}{\max\{U_j^1, U_j^2, \cdots, U_j^n\}}.
\]

Here, for the mutation, the system creates a new layout based on the relationships the system estimated. Considering the most-related section of each section, mutually-related sections get closer in the new layout so that a viewer can easily find out important sections which are related to the section he is looking at.

For crossover, the system compares each value in the same sections in the two selected content designs. The possibility of output values is based on the attracting evaluation and unconvincing evaluation of each content design so
that attracting and convincing sections are selected for the crossover. For example, when the system crossovers two content designs, \( a \) and \( b \), in the section \( k \), the probability that graphical property in each content design is used is in the ratio of \( E^{a,b} \) to \( E^{b,k} \).

5. Results and Evaluation

5.1 Accuracy of Behavior Recognition

Firstly, we evaluated the accuracy of behavior recognition through Microsoft Kinect\( ^\text{TM} \) from two aspects: accuracy of viewers’ gesture recognition and gaze estimation. We asked a subject to perform one indicated gesture in front of the public display, and also asked him to look at one indicated section. We did the evaluation a hundred times, and as the result, the system recognized the gesture correctly in 91% of the tests, which shows the validity of the gesture recognition. Additionally, the system recognized the gaze point correctly in 88% of the tests. This tells us the sufficient accuracy of gaze estimation, while the accuracy rate still can be improved utilizing eye movements through the further development of the resolution of the capture device.

5.2 Evaluation of Evolved Content Designs

In order to evaluate the proposed system empirically, we made up some sets of three content designs: (a) one from the seventh generation where \( N \) is six and the time interval of each content design is ten minutes; (b) one from the first generation; and (c) manually made content design. In this case, five to ten Japanese computer science students who are not professional in the field of public display attended as viewers in each content design. The reason why \( N \) is so small is that content designs of academic conference posters are confined within a few variety of designs and there could be similar designs even \( N \) was large. Figure 7 shows example results of (a), (b) and (c).

In order to compare attractiveness of design, convincingness of content and reflection of mutual relationship of sections to layout in (a), (b) and (c), we showed these produced content designs to twenty-five Japanese subjects. As was the case with the viewers, these subjects were general computer science students, who have an ability to understand the content. As we mentioned above, there are different sets of (a), (b) and (c), and thus these people did not always see the same set of (a), (b) and (c). After this, we asked them how they feel about the three content designs.

The results in Table 2 show that, in every aspect, about 70-85% of the subjects thought evolved content designs are better than the initial content designs. This suggests most of the content designs got improved through the system. In addition to this, in every aspect, more than half of the subjects thought evolved content designs are better than the manually-designed one.

What is interesting is that the number of subjects who thought (a) is better than (b) and the number of subjects who thought (a) is better than (c) were almost the same in terms of “Attractiveness”. In particular, the attractiveness of manual design especially reflects the sender’s ability, and who created the design in this test was not a professional designer. This can be a reason why the numbers are almost the same, and this result suggests that even randomly generated design can be as attractive as non-professional manual design, only if focusing on the appearance. This makes us re-acknowledge how it is difficult to seek for attractive design. However, since about 70% of subjects thought (a) is better than (b) and (c), it can be argued that the system did improve the attractiveness, which validates the effectiveness of our system.

On the other hand, as for “Layout of sections”, the difference between “(a) is better than (b)” and “(a) is better than (c)” is the largest, and while about 85% of subjects thought (a) is better than (b), only about 50% of subjects thought (a) is better than (c). This means while layouts were improved through the evolutionary method well, the improved layout did not exceed the manual layout in most cases. One possible reason is that layouts of conference poster are limited in some styles. Because of this, there are not so many varieties of layout, and in some case the evolved layout and the manual layout look similar. We need to seek the potential of evolved layout in more general usage, where various layouts can be considered.

6. Discussion

The result shows the acceptable effectiveness of our system, while there are some allowances for extensions. Most importantly, viewers’ mood can be considered from other contexts. One point is interactions among “viewers”. In this system, we regarded viewers as individual people. However, when viewers are looking at public display, in many cases, they interact with each other: inviting others to look at the content, discussing about the content and so on. We need to consider the “mutual influences among individuals”
for more understanding of on-site mood, as hardware technologies develop.

Another issue is to consider individuality. In our system, there is an assumption that there are a reasonable number of people who have similar tastes, especially in the case the public display is set at a specific location such as a conference, and the result shows pragmatical effectiveness. However, when there is a small number of people, or when we extend the method for more generalized applications, our system can be made more efficient by considering the difference of individual viewers, through the use of database managing individual differences and modifying the content design based on the individual viewer’s preference. Additionally, by combing other areas of information engineering, the system can be extended. For example, with natural language analysis, the system can automatically estimate and analyze the amounts of information of the sections. Also, if the system can clearly understand whether viewers’ gestures are caused by its content or graphical properties, the system can modify the content design in more effective way, as we can deal with the semantics of behaviors more plausibly.

7. Conclusion

In this paper, we presented a novel pragmatic system for surveying, evaluating and modifying the conference poster content design. This method realizes the public display system which automatically learns what structure and content are attractive and makes the content more adapted to the viewers, through utilizing the viewers’ involuntary behaviors, including their gestures and movements, and the idea of GA. The system takes away the current gap which lies between the viewers’ expectations and the content actually displayed on public display. We proved the feasibility of the method empirically through the development of a Mood-Learning Public Display for displaying academic conference posters.

Acknowledgments

The work is supported in part by a Grant-in-Aid for the Leading Graduate School Program for “Science for Development of Super Mature Society” from the Ministry of Education, Culture, Sports, Science and Technology in Japan.

References

[1] K. Nagao and I. Fujishiro, “Making digital signage adaptive through a genetic algorithm — utilizing viewers’ involuntary behaviors,” Proc. INSTICC International Conference on Computer Vision Theory and Applications, vol.2, pp.54–59, 2013.
[2] K. Nagao and I. Fujishiro, “Improved structured content design of digital signage evolutionarily through utilizing viewers’ involuntary behaviors,” Proc. Cyberworlds 2013, pp.108–115, 2013.
[3] K. Nagao and I. Fujishiro, “Understanding viewers’ involuntary behaviors for adaptive digital signage,” Proc. ACM Symposium on Applied Perception, p.145, 2013.
[4] F. Alt, S. Schneegaß, A. Schmidt, J. Müller, and N. Memarovic, “How to evaluate public displays,” Proc. 2012 International Symposium on Pervasive Displays, pp.17:1–17:6, 2012.
[5] S. Gotta, J.R. Huang, P. Jonathon, and H. Wechsler, “Mixture of experts for classification of gender, ethnic origin, and pose of human faces,” IEEE Trans. Neural Netw., vol.11, no.4, pp.948–960, 2000.
[6] M. Pantic and L.J.M. Rothkrantz, “Automatic analysis of facial expressions: The state of the art,” IEEE Trans. Pattern Anal. Mach. Intell., vol.22, no.12, pp.1424–1445, 2000.
[7] G. Beyer, F. Alt, J. Müller, A. Schmidt, K. Isakovic, S. Klose, M. Schieve and I. Haulsen, “Audience behavior around large interactive cylindrical screens,” Proc. SIGCHI Conference on Human Factors in Computing Systems, pp.1021–1030, 2011.
[8] J. Müller, R. Walter, G. Bailly, M. Nisch and F. Alt, “Looking glass: A field study on noticing interactivity of a shop window,” Proc. SIGCHI Conference on Human Factors in Computing Systems, pp.297–306, 2012.
[9] D. Vogel and R. Balakrishnan, “Interactive public ambient displays: Transitioning from implicit to explicit, public to personal, interaction with multiple users,” Proc. 17th Annual ACM Symposium on User Interface Software and Technology, pp.137–146, 2004.
[10] G. Schiavo, E. Mencarini, K.B. Vovard, and M. Zancanaro, “Sensing and reacting to users’ interest: An adaptive public display,” CHI ’13 Extended Abstracts on Human Factors in Computing Systems, pp.1545–1550, 2013.
[11] T. Ojala, V. Kostakos, H. Kukka, T. Heikkinen, T. Linden, M. Jurmu, S. Hosio, F. Kruger, and D. Zanni, “Multipurpose interactive public displays in the wild: Three years later,” Computer, vol.45, no.5, pp.42–49, 2012.
[12] F. Alt, A.S. Shirazi, T. Kubizra, and A. Schmidt, “Interaction techniques for creating and exchanging content with public displays,” Proc. SIGCHI Conference on Human Factors in Computing Systems, pp.1709–1718, 2013.
[13] R. Ballagas, M. Rohs, and J.G. Sheridan, “Sweep and point and shoot: Phonecam-based interactions for large public displays,” CHI ’05 Extended Abstracts on Human Factors in Computing Systems, pp.1200–1203, 2005.
[14] N. Wouters, J. Huyge, and A.V. Moere, “Openwindow: Citizen-controlled content on public displays,” Proc. 2nd ACM International Symposium on Pervasive Displays, pp.121–126, 2013.
[15] G. Schiavo, M. Milano, J. Saldivar, T. Nasir, M. Zancanaro, and G. Converto, “Agoraz2.0: Enhancing civic participation through a public display,” Proc. 6th International Conference on Communities and Technologies, pp.46–54, 2013.
[16] R.A. Clark, Y.H. Pua, K. Fortin, C. Ritchie, K.E. Webster, L. Denehy, and A.L. Bryant, “Validity of the microsoft kinect for assessment of postural control,” Gait and Posture, vol.36, no.3, pp.372–377, 2012.
[17] A. Khan, J. Matejka, G. Fitzmaurice, and G. Kurtenbach, “Spotlight: Directing users’ attention on large displays,” Proc. SIGCHI Conference on Human Factors in Computing Systems, pp.791–798, 2005.
[18] N. Singh and S. Bhattacharya, “A GA-based approach to improve web page aesthetics,” Proc. First International Conference on Intelligent Interactive Technologies and Multimedia, pp.29–32, 2010.
[19] H. Gunes and M. Piccardi, “A bimodal face and body gesture database for automatic analysis of human nonverbal affective behavior,” Proc. 18th International Conference on Pattern Recognition, vol.1, pp.1148–1153, 2006.
Ken Nagao received his bachelor degree at Keio University in 2012, and received his Master of Science in Engineering at Keio University in 2014. He is currently pursuing his second master degree in business and commerce as a partial fulfillment of Leading Graduate School Program for "Science for Development of Super Mature Society" from the Ministry of Education, Culture, Sports, Science and Technology in Japan.

Issei Fujishiro received his Master of Engineering in information sciences and electronics from University of Tsukuba in 1985 and his Doctor of Science in information sciences from The University of Tokyo in 1988. He is currently a professor of information and computer science in Keio University. His major research interests include volume graphics, visualization lifecycle management, and multi-modal information display. He is currently serving as a vice president of SAS and the chair of SIG on Visual Computing of IIEEJ.