SPAL: A Sensor of Physical-World Attention using Laser Scanning

Yasunori Yakiyama† Oru Mihirogi† Niwat Thepvilojanapong†‡ Kazunori Umeda§‡ Yoshito Tobe†‡
†Tokyo Denki University, Japan §Chuo University, Japan ‡CREST, JST, Japan
E-mail: {yakly, oru, wat}@u-netlab.jp, umeda@mech.chuo-u.ac.jp, yoshito_tobe@osoite.jp

Abstract

Although human activities in the World Wide Web are increasing rapidly, we still need to identify how things such as pictures in a museum are received attention in the real world. To measure people's attention, we propose SPAL, a Sensor of Physical-world Attention using Laser scanning. A key scheme to identify the attention is determining how long a moving person lingers around the target object. In this paper, we describe the design of SPAL and show the results from the real-world experiments.

1 Introduction

Laser scanners and image sensors have been used for measuring human behavior [10]. Most of previous works focus on the detection or tracking of human. The measurements of such works are applied to traffic management and security systems. However, a measurement of people's attention has not been studied sufficiently. We believe that a measurement of people's attention benefits human society as a metric for the purpose of comfortable urban activity. For example, we can place objects at the right places in an exhibition hall or a museum according to people's attention.

This paper proposes a *SPAL (Sensor of Physical-world Attention using Laser scanning)* system to calculate a degree of people's attention from their behavior in the real world. Our system focuses on attention of target objects in a conventional hall and museum. In the first step, the system detects human objects based on data measured from laser scanners. It then determines staying people of each detected human object, and calculates a degree of people's attention as an output of the system. A *Degree of Real-World Attention (DRWA)* is defined as a metric to show a degree of people's attention. We propose three measurement models to calculate the DRWA.

We performed the experiments to study the performance of SPAL system. The experiments were done indoor by placing three posters as target objects. The DRWAs are calculated and compared with the actual attentions of people which are taken from questionnaires.

The rest of this paper is organized as follows. Section 2 discusses related work. Section 3 proposes three models to calculate physical-world attention. We describe the system architecture in Section 4 and evaluate the system in Section 5. We conclude our study in Section 6.

2 Previous work

Attention is often referred to *visual attention* in image processing when recognizing objects. Ma et al. [4] defines a user attention model which estimates attentions that viewers may pay to video contents. With the same purpose, many visual attention models have been proposed for video summarization [2, 6, 5]. In addition, a human-robot interaction system requires attention mechanisms to comprehend a situation [3]. Our scope is different from the above visual attention studies. We define *attention* as interest of people in the real world.

Our system exploits techniques of people tracking, pedestrian counting, and crowd flow detection as fundamental basis. Zhao et al. proposed to scan feet of pedestrians by laser scanners and analyzed walking trajectory based on a pedestrian model [11, 10]. They also applied their system to visualize passenger flow in railway stations [7]. In intelligent transportation system, a system to recognize pedestrians provides useful information for drivers in order to achieve safe driving [1]. A stereo camera is also used to recognize pedestrians from images [9, 12]. Previous works use laser scanners and stereo cameras as two major devices for tracking and detecting flow of people. Although a stereo camera can be used to measure a distance from an object, it is affected by environment such as light and has a privacy issue in public space. Thereby, we consider using laser scanners in our work.

3 Models for Real-World Attention

Since the SPAL calculates real-world attention from a state of people staying, we define *Value of People Staying (VPS)* as the number of people who stay nearby a target object. Let U be a set of N target objects. We define *Degree of Real-World Attention (DRWA)* of the target object x ($x \in U$) as a ratio of the attention level of the target object x and that of all target objects in U (Equation 1).

$$DRWA(x) = \frac{VPS(x)}{\sum_{i=1}^{N} VPS(x_i)}. \quad (1)$$

3.1 Counting states of People Staying (CPS) Model

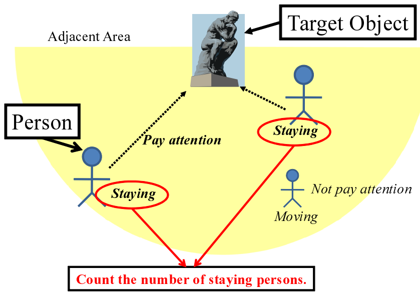


Figure 1: The measurement method of CPS model.

The *Counting states of People Staying (CPS)* model determines the state that people stay at a certain point more than a threshold period as a state of paying attention to a target object (Figure 1). Let $CPS(x,t)$ be the number of staying people at time t for the target object x.

$$CPS(x,t) = \sum_{i=1}^{DP(x,t)} f(x,t,i), \text{ where} \quad (2)$$

$DP(x,t)$ is the number of detected people at time t and $f(x,t,i)$ is the function to decide whether people i is staying at time t. In particular, $f(x,t,i)$ is one if staying period is longer than a threshold, otherwise zero. The VPS of the target object x is a sum of $CPS(x,t)$ for measurement period T, i.e., $\sum_{t=0}^{T} CPS(x,t)$.

3.2 Distance based Weighted Counting states of People Staying (DWCPS) Model

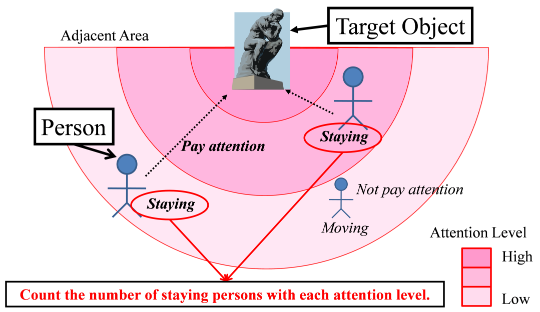


Figure 2: The measurement method of DWCPS model.

In the CPS model, all staying people are counted as the same value, namely, one. However, in an actual situation, it is considered that the nearer the distance between people and a target object, the higher the attention level of people is (Figure 2). Therefore, we define *Distance based Weighted Counting states of People Staying (DWCPS)* model that includes a weighted parameter corresponding to the distance between people and the target object when calculating the VPS.

$$DWCPS(x,t) = \sum_{i=1}^{DP(x,t)} w(x,t,i) f(x,t,i). \quad (3)$$

$$w(x,t,i) = \begin{cases} 1, & d(x,t,i) \leq d_1 \\ \dfrac{d_0 - d(x,t,i)}{d_0 - d_1}, & d_1 < d(x,t,i) < d_0 \\ 0, & d_0 \leq d(x,t,i). \end{cases} \quad (4)$$

$w(x,t,i)$ is the weighted parameter at time t of people i for the target object x, and $d(x,t,i)$ is the distance between people i and the target object x at time t. The VPS of DWCPS model is a sum of $DWCPS(x,t)$ for measurement period T, i.e., $\sum_{t=0}^{T} DWCPS(x,t)$.

3.3 Counting Grids existed states of People Staying (CGPS) Model

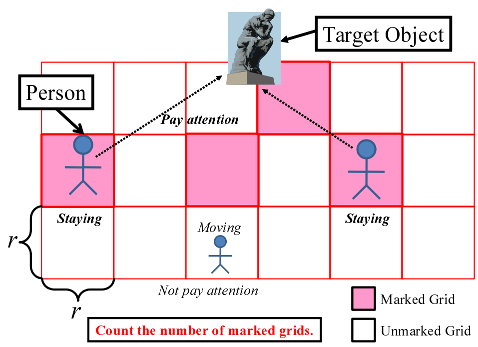


Figure 3: The measurement method of CGPS model.

We propose *Counting Grids existed states of People Staying (CGPS)* model which divides a scanning area into multiple grids (Figure 3). Let $G(X_i, Y_i)$ be a state of a 2D grid, where X_i and Y_i indicate indices of the grid. In an initialization step, all grids are unmarked and the values are zeros, i.e., $\forall G(X_i, Y_i) \leftarrow 0$, when $t = 0$. If people move into an unmarked grid and stay longer than a threshold, the grid is marked and the state is changed to one. Once a grid is marked, the state is permanent until the end of the measurement. The VPS is the number of marked grids at $t = T$, i.e., $\sum_{i=0}^{\forall X} \sum_{j=0}^{\forall Y} G(X_i, Y_j)$.

4 System Architecture and Implementation

4.1 System Architecture

This section briefly describes the processing flow of SPAL system. First, the system performs a step of detecting human objects by using a Human-Detection algorithm. The algorithm is as follows: (1) generate background data by scanning an area of interest while no one exists, (2) compute a margin between current data and background data in counterclockwise direction, (3) if the margin is longer than a threshold (5 cm), the current point of current data is set as a begin edge, (4) continue until the margin is shorter than the threshold where an end edge is found, (5) if the distance between the begin and end edge satisfies the condition (between 30 cm and 70 cm), the recognized object is determined as a person.

The system then uses a Staying-Detection algorithm to determine whether a human object is a staying person. The algorithm is as follows: (1) store location of detected person for five cycles (total 1.065 seconds), (2) compute the vibration of the people which is the difference between the maximum and minimum values in five cycles, (3) if the vibration is less than a threshold (30 cm), the person is recognized to be a staying one. The system logs positions of staying and moving people. Then the VPSs are counted for each target object individually, and the DRWAs are calculated by using Equation 1.

4.2 Prototype Implementation

The SPAL system consists of a laser scanner which connects to a processing node through RS-232C interface. The laser scanner is LMS-200 developed by SICK in Germany [8]. The laser is 905-nm near infrared rays and the safety class is 1A. The laser scanner scans counter-clockwise direction with maximum scanning angle of 180 degree. The angular resolution is 0.5 degree which means the laser scanner has 361 scanning steps. The maximum scanning distance is 80 m and the distance resolution is 1 cm. In the SPAL system, we place the laser scanner 140 cm above the ground level. Thereby the scanning plane is approximately the chest of an adult. The scanning rate is 4.7 Hz.

The processing node is a laptop computer running Windows Vista, and we use .NET Framework 2.0 as runtime environment. We developed an application software and installed in the processing node. The application software obtains scanned data from the laser scanner, and then analyzes the data by detecting a person and calculating the degree of real-world attention. GUI is also developed for easy usage.

5 Performance Evaluation

5.1 Experiment Setup

We conducted the experiments by placing three posters as target objects. The posters contain different contents. In particular, the laser scanners are placed in front of the target objects at a distance of 200 cm. The scanning angle is between 30°and 150°, and the scanning area is the area within 300-cm radius from the laser scanners.

All testees observed the posters freely and evaluated their interests according to the five-score system in a questionnaire: very high (5), high (4), normal (3), low (2), and very low (1). During the experiments, we logged distance data scanned by the laser scanners. The experiments had been conducted for 10 minutes, and the number of testees was six people. After the experiments, we analyzed logged data and compared attentions calculated from the proposed models with the testees' actual attention obtained from the questionnaires.

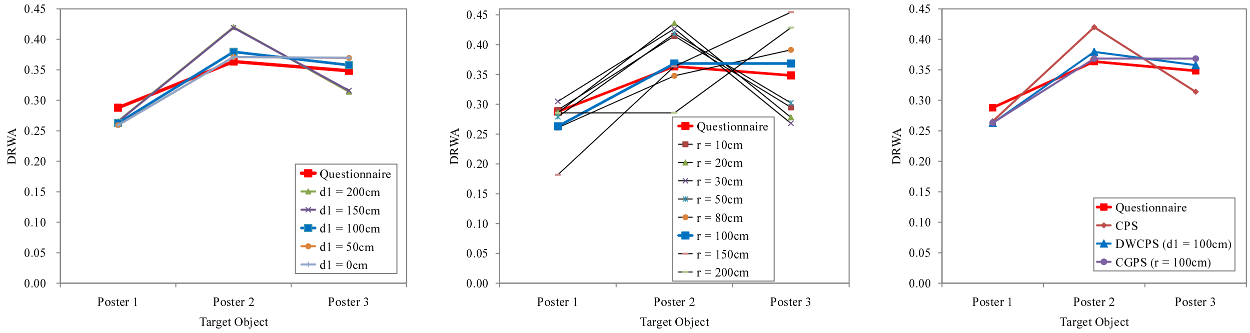
For the actual attention, the scores (1-5) taken from the questionnaires are summed up for each target object. The actual DRWA is a ratio of a cumulative score of a target object and that of three target objects. The DRWAs based on scanned data are calculated by following the three proposed models. For the DWCPS model, the value of d_0 is set to 200 cm, while d_1 is set to 0, 50, 100, 150, and 200 cm. Also the value of r in the CGPS model is set to 10, 20, 30, 50, 80, 100, 150, and 200 cm.

5.2 Experimental Results

First, we compare the results of the DWCPS model with those of questionnaires in Figure 4(a). The results of d_1 = 0, 50, and 100 cm highly match those of questionnaires, i.e., the errors are below 10%. Although the results of d_1 = 150 and 200 cm are not so good as those of d_1 = 0, 50, and 100 cm, their errors are lower than 15%. We conclude that the DWCPS model is a good candidate for determining people's attention which correlates to a distance between a person and a target object as expressed in Equation 3. The results also show that the precision of the model drops as the value of d_1 increases beyond half of the observation range (200 cm) because the weight of people staying far from a target object is too high.

Next, we compare the results of the CGPS model with those of questionnaires in Figure 4(b). The results of r = 100 cm are the best match with those of questionnaires and the errors are below 10%. The errors are less than 23% if r is less than 80 cm. As one would expect, large grid size (r = 150 cm) has low precision, i.e., the errors increase to 37% because large grid is likely to judge moving people as staying people. The results show that the CGPS model has high precision in determining people's attention and the grid size should be 100 cm.

Finally, we compare the results of three measurement models: CPS model, the best result of DWCPS model (d_1 = 100 cm), and the best one of CGPS model (r = 100 cm). The results of questionnaires are also included as references and the comparison is shown in Figure 4(c). The errors of



(a) People's attention of DWCPS model.

(b) People's attention of CGPS model.

(c) Comparison of three mesurement models.

Figure 4: Experimental results.

the CPS model are higher than those of DWCPS and CGPS models, i.e., the error are 15% in the worst case (poster 2). When people pay attention to a target object, they may change their position in order to see the target object from different views. The CPS model cannot capture such people's attention correctly. The average errors of the DWCPS and CGPS models are approximately the same (5%). The parameters of both models (d_1 and r) affect the calculation of people's attention. Therefore, we have tested multiple values of both parameters and the results show the appropriate values as discussed above.

6 Conclusion

This paper has studied the tendency of people's attention in the real world by using the proposed SPAL system. The system calculates people's attention based on three proposed measurement models, i.e., CPS, DWCPS, and CPGS models. We have implemented the system and conducted the experiments. Raw data obtained from laser scanners are input of the system, and people's attention based on three models has been calculated to study the performance of the system. The results show that the proposed models are good candidates for determining people's attention which correlates to a distance between people and a target object. we note here that a problem of personal privacy does not concern the SPAL system because our system does not recognize an individual.

References

[1] C. Curio, J. Edelbrunner, T. Kalinke, C. Tzomakeas, and W. Seelen. Walking pedestrian recognition. *IEEE Trans. Intelligent Transportation Systems*, 1(3):155–163, 2000.

[2] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(11):1254–1259, Nov. 1998.

[3] I. M, O. T, and I. H. Physical relation and expression: joint attention for human-robot interaction. *IEEE Trans. Industrial Electronics*, 50(4):636–643, Aug. 2003.

[4] Y.-F. Ma et al. A generic framework of user attention model and its application in video summarization. *IEEE Trans. Multimedia*, 7(5):907–919, Oct. 2005.

[5] Y.-F. Ma, L. Lu, H.-J. Zhang, and M. Li. A user attention model for video summarization. In *Proc. 10th ACM Int. Conf. Multimedia*, pages 533–542, Dec. 2002.

[6] Y.-F. Ma and H.-J. Zhang. A model of motion attention for video skimming. In *Proc. IEEE Int. Conf. Image Processing*, pages 129–132, Sept. 2002.

[7] K. Nakamura et al. Visualizing passenger flow in railway station using laser scanners. In *International Conference on Computers in Urban Management and Urban Planning*, 2005.

[8] SICK. Laser measurement systems. http://www.sick.com/home/factory/catalogues/auto/lmsindoor/en.html.

[9] J. C. Wohler, J. Anlauf, T. Portner, and U. Franke. A time delay neural network algorithm for real-time pedestrian recognition. In *Proc. Intelligent Vehicles*, pages 247–252, 1998.

[10] H. Zhao et al. A laser scanner system for acquiring walking trajectory data and its possible applications to behavioral science. In *Studies in Human and Social Sciences with GIS: Chapter 3*, Taylor & Francis/CRC Press, Nov. 2005.

[11] H. Zhao and R. Shibasaki. Pedestrian tracking using multiple laser range scanners. In *International Conference on Computers in Urban Planning and Urban Management*, 2003.

[12] L. Zhao and C. Thorpe. Stereo and neural network-based pedestrian detection. In *Proc. ITSC*, pages 289–303, 1999.