

# Delivering Active Fashion Advertisement to Potential Viewers

Mitsuru Nakazawa<sup>(✉)</sup>, Jiu Xu, and Soh Masuko

Rakuten Institute of Technology, Rakuten, Inc., Rakuten Crimson House,  
1-14-1 Tamagawa, Setagaya-ku, Tokyo, Japan  
{mitsuru.nakazawa, jiu.xu, so.masuko}@rakuten.com

**Abstract.** Nowadays, more and more fashion stores have started to introduce digital signage displays for dynamically switching advertisement contents of their products. To attract people's attention and interest towards a display, some signage systems distribute suitable contents for each individual based on personal attributes, acquired through a camera. However, if there are people who are not interested in the digital signage around the display, suitable contents are not necessarily distributed to those who really intend to view it. Therefore, in this paper we propose a digital signage that can take into account people who have higher potential of viewing the display. First, we acquire human behavior and personal attributes of each person. Then, the potential degree of viewing the display is calculated for each person. Next, our system searches content candidates for each person based on their personal attributes. Finally, the distributed contents are determined by selecting more candidates for people with a higher viewing potential.

**Keywords:** Digital signage · Fashion advertisement · Viewing potential

## 1 Introduction

Generally, in fashion stores, advertisements come in the form of actual products or their posters. Recently, more and more stores have begun introducing digital signage displays as the advertisement contents can be easily updated. At the end of 2016, over 37 million digital screens were in use all over the world [1].

To change the advertisement contents of a digital signage, one of the simplest ways is to loop them on repeat. To make people more attracted, some signages can change contents by the time of day and using weather information [8], or based on social media [3], or a specific human action as the trigger such as approaching the display [2]. Moreover, to provide suitable contents for a certain person, that is, "for content personalization", [5] can change contents based on gender and age as personal attributes, which are acquired using a camera mounted on the display. However, if there are people who are not interested in the digital signage around the display, suitable contents are not necessarily distributed to those who really intend to view it.

To solve the aforementioned issue, in this paper we propose a digital signage system that can take people who have higher potential of viewing the display into account, as shown in Fig. 1. In our system we first use an RGB-D sensor to acquire human behavior, including walking speed and face orientation, together with personal attributes such as gender, and the representative colors of the upper and the lower body. Then, based on their behavior, the potential degree of viewing the display is calculated for each person. Next, our system searches content candidates for each person based on their personal attributes. Finally, the distributed contents are determined by selecting more candidates for people with a higher viewing potential.

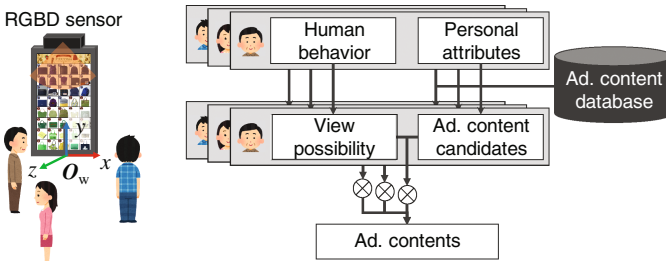


Fig. 1. System overview

## 2 Proposed System

Our proposed system assumes that people are passing in front of the system as shown in Fig. 1. To acquire human behavior and personal attributes, a sensor is mounted on the top of the system; we chose to adopt an RGB-D sensor as it can easily detect people and obtain point cloud data of each person in the metric scale. The angle of the sensor is set downwards so that it can observe people from a bird's-eye view. The following subsections describe the details of our proposed system, which observes time-series point cloud data from the RGB-D sensor as source data.

### 2.1 Acquiring Human Behavior and Personal Attributes

As preparation for the acquisition, we extract time-series point cloud data of each person from the source data. First, the coordinate system is converted from the camera coordinates to world coordinates, where the  $xy$  and the  $xz$  plane are aligned to the display and the floor plane respectively and the origin  $O_w$  is located at the bottom center of our system as shown in Fig. 1. Then, the 3D region of each person is detected at each frame by Munaro et al.'s method [4] focusing on the shape of a human head. Finally, person tracking is performed by matching detected persons between the current and the previous frame. In our system, a greedy algorithm is used to find matching correspondences that

minimize the Euclidean distance between the center position of a person at the current frame and the estimated position from a person’s walking trajectory.

From the time-series point cloud data of each person, we obtain their behavioral and personal attributes. In terms of behavioral attributes, we acquire the position  $\mathbf{x}_i$ , the vector of walking speed  $\mathbf{v}_i$  and horizontal face orientation  $\theta_i$  of the  $i$ -th person. In this paper,  $\theta_i$  is set as 1 if Viola and Jones’ detector [7] detects a frontal face from the color image of his or her head region; otherwise 0. As personal attributes, we acquire the representative colors of the upper and lower body, and the gender of each person. Additionally, a compatible color with the two representative body colors is chosen based on the accent color and separation color of fashion color theory [6].

### 2.2 Calculating Potential Degree of Viewing the Display

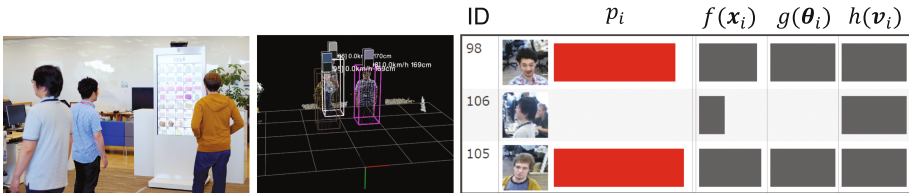
The potential degree of viewing the display of the  $i$ -th person  $p_i$  is calculated from his or her personal behavior. In the calculation, we consider “from where, how frequently, and how they are walking whilst viewing” as the state of viewing behavior.

$$p_i(\mathbf{x}_i, \theta_i, \mathbf{v}_i) = f(\mathbf{x}_i)g(\theta_i)h(\mathbf{v}_i), \tag{1}$$

where  $\theta_i$  is the vector of face orientation in the past  $N$  frames;  $f$  is a negative piecewise linear activation function of the distance between the person and the display plane, which returns higher values for people closer to the display;  $g$  is a monotone increasing function of the frequency of detected frontal faces in the  $N$  frames, which returns higher values for people viewing the display more frequently; and  $h$  is a step function of the walking speed that returns 0 for a person walking more quickly than a threshold speed, otherwise 1.

### 2.3 Distributing Suitable Advertisement Contents

In preparation for the content distribution, a database is constructed beforehand so that suitable advertisement contents can be searched by personal attributes. Our database consists of the image, gender information (i.e. men’s, ladies’ and unisex), the representative color and the type of clothes (e.g. tops and bottoms) of each fashion item.



**Fig. 2.** Developed system (Left: scene image, middle: analyzed point cloud data, right: potential degree  $p_i$  and its component parameters)

To distribute suitable advertisement contents, first our system searches content candidates for each person from the database by their gender and compatible color. Then, the distributed contents are selected from the content candidates of each person so that the ratio of the number selected contents for each person is the same as the normalized potential degree  $p'_i = p_i / \sum_i p_i$ . The above procedure is repeated to update advertisement contents at a certain interval.

### 3 Conclusion

In this paper, we proposed a digital signage system that can take people who have higher potential of viewing the display into account. First, we acquired the human behavior and the personal attributes from time-series point cloud data observed from a RGB-D sensor. Next, the potential degree was calculated from the human behavior with consideration of their state of viewing behavior. Then, content candidates that are suitable for each person are searched from the database by using the personal attributes. Finally, distributed contents are selected from the candidates so that the ratio of selected content candidates for each person matches their normalized potential viewing degree.

As shown in Fig. 2, we have finished developing the proposed system, which will be demonstrated in the interactive session of ICEC2017. As future work, for better personalization of advertisement contents, we will try to acquire new personal attributes including age or the fashion style of pedestrians from the RGB-D sensor.

### References

1. Digital Signage M2M and IoT Applications - 2nd Edition. Research and Markets (2017)
2. ImpactTV: Pista (in Japanese). <http://impacttv.co.jp/product/>. Accessed 10 May 2017
3. Mevato: Social marketing display for any screen. <http://www.mevato.com/display>. Accessed 10 May 2017
4. Munaro, M., Basso, F., Menegatti, E.: Tracking people within groups with RGB-D data. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (2012)
5. Anewtech Systems: Intelli-signage: age and gender recognition system. <http://intelli-signage.com/technologies/age-and-gender-recognition-system-agrs/>. Accessed 10 May 2017
6. Tsurumaki, A.: How to coordinate fashion colors (in Japanese). <https://tokila.jp/how-to-coordinate-colors/>. Accessed 10 May 2017
7. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. I-511–I-518 (2001)
8. Ads of the world: Mcdonald's McMuffin sunrise. [http://adsoftheworld.com/media/outdoor/mcdonalds\\_mcmuffin\\_sunrise](http://adsoftheworld.com/media/outdoor/mcdonalds_mcmuffin_sunrise). Accessed 10 May 2017