Poster: WatchYouWatch - A Web-Cam Based Natural Customer Attention Tracking Shelf

Mengxin Cao
Beijing University of Posts and Telecommunications

Zhiyuan Liu
Tsinghua University
Haidian District, Beijing, 100084 P.R. China

Haotian Long
Beijing Jiaotong University
Beijing, P.R. China

Guang Li
Megvii Inc. (Face++)
Beijing, P.R. China
liguang@megvii.com

Abstract

In this paper, we introduced WatchYouWatch, a web-cam based smart shelf that can directly capture customer's attention of the products on the shelf. By implicitly deploying ordinary web cameras on the rack, we could capture the gaze points of customers on the products scattered in different areas and further deduct which one is more popular. Such approach builds upon the well studied eye-tracking technology and interacts with users naturally. Based on this idea, we implemented a prototype customer attention tracking shelf, integrating web-cams, eye-tracking, and attention monitoring. A pilot study with 10 volunteers demonstrated the practicability of our proposed system.

1 Introduction

Below the line user experiment is the main source of obtaining the behavior data of customers. Various product arrangement plans may be proposed in the assessment of newly designed products and testers are usually recruited to participate in the investigation. Currently, the most common approaches to collect customer preferences over different products are surveys and interviews, which suffers from high labor cost and common questionnaire design issues such as rating level inconsistencies, ambiguous or unintelligible questions, etc. Moreover, such ineffectiveness usually leads to a cognitively biased result [6] and a simple, natural yet credible method to determine user preferences is of vital importance.

As the state of the art, eye-tracking technology has been widely applied in fields such as virtual reality, human-computer interaction and market research [7]. The advancement in computer vision based on deep learning techniques further boosted the development of eye-tracking researches. Even with built-in cameras, images can be understood by machines and achieve user's eye tracking without calibration. Recent studies have shown eye-tracking could help us better carry out market research and analyze customer's attentions. Such as customer's attention to web advertisement or TV commercials [5], visual influence on in-store buying decisions[2], attention analysis on mutated brand names [1], digital product package analysis using eye-tracking [4].

To directly track customer's attentions in product picking and thus evaluate the on-shelf products, we propose to use eye-tracking based technologies to measure customer's attentions. We could setup a web-cam on the rack that is concealed to customers to collect their real time attentions to various products, which can be used to aid the analysis of user preferences over these products. Our proposed approach repacks the ubiquitous shelves with minimal modification cost and highly flexible in design. Customers interact with the system spontaneously and directly provide authentic attention input almost unconsciously. Such smart shelf solution can be widely deployed in supermarkets, vending machines or shopping malls, which characterizes high market potential.

2 Implementation

We believe that some basic design rationale should be met so as to better monitor customer's authentic attention. First, the physical appearance of the proposed design should not diverge so much compared to present prevalent shelves so that it is acceptable for users. The system requires web-cams and processing nodes but they should be installed in a covert manner to reduce the possible interference on customer's buying behaviors. In addition, considering the application scenario and huge market needs, the system deployment costs should be as low as possible in principle.

Based on the analysis above and constraints, our approach is to repack the ordinary shelves. Web-cams and processing units are embedded behind the rack to conceal them from customers. Meanwhile, we have the web-cams connected to processing nodes, providing vision information such as gaze points about users afront of the shelf. After meaningful information extracted from the raw frames of the video stream, the system then identifies the product that received greater attention and reports the decision from this shelf to the system's back-end server, which aggregates the results from all nodes and ranks products from most popular to the least.

To prove the feasibility of our proposed approach, we constructed a repacked prototype based on ordinary shelves.

International Conference on Embedded Wireless Systems and Networks (EWSN) 2019
25–27 February, Beijing, China

ISBN: 978-0-9949886-3-8

Figure 1. Overview of WatchYouWatch.

As illustrated in Fig.1, the combination of web camera (Logitech C270 web-cam), Raspberry Pi (as the processing end node) and LEDs (to demonstrate customer's current attention area and should be removed in deployment) was setup as a bundle adjunct to the shelf. Each tier of the shelf is divided into separate areas and only one product is placed at each area. Testers attempt to find their most interested product by naturally looking at the displayed objects, which is enough for the system to track their attention movements and measure the popularity of each product. The product replacement frequency is determined by the relative popularity among products and the frequency increases for the least popular ones.

The user-shelf interaction remains straightforward and intuitive. The embedded web-cam captures users' gaze point when they are picking products afront the shelf. Our prototype used the existing eye-tracking API provided by Face++ [3], and Raspberry Pi (RPi) is responsible for tracking users' attention duration and movements as well as reporting to the back-end server to complete the analysis of user preferences. The interaction described above happens in real time, behind the scene and users are unconscious about what have happened.

We recruited 10 volunteers to participate in our pilot study, which consists of 7 males and 3 females, 5 of them with glasses. Two groups of experiments conducted, and the volunteers took turns to pick their preferred product. Group A consists of two different liquid laundry detergent while group B is three different soft drinks. For each experiment, the corresponding LED will be lightened when customers look at that product as an auxiliary indication of customer's current attention. Meanwhile, the volunteers are also required to write down their preferences after the selection process in order to compare the effectiveness of our system. The experiment took 1 hour and 90% of the products in group A correctly matched with user's preferences and 80% for group B, which demonstrates the practicability of our system.

3 Discussion and Future Work

The usability our proposed system is effective in measuring the attention received by products on the shelf. Yet several problems still exist in current prototype. First, the equipment costs are quite high compared to infrared sensor based or ToF camera-based methods. However, such cost difference can be balanced by the low repack costs, high flexibility and better performance provided by our system. Infrared sensors are sensitive to distances and hence result in nonnegligible variances. Single web-cam can be deployed to monitor multiple areas and provide high dimensional customer attributes such as age, gender, with glasses or not, etc. In addition, current eye-tracking API service suffers from users' subjective factors such as height, eye size, and with glasses or not. Products are scattered on the shelf and the space is poorly utilized in our prototype.

Since we have only constructed a prototype, much more needs to be done to transform our system into a production ready service. First, the layout design of the web-cams embedded in the shelf should be optimized to conceal them as much as possible to enable more natural interaction between users and our system. Meanwhile, the system costs should be minimized either by improving the capability of single camera or deploying with cheaper cameras. To achieve more robust tracking performance and recognition accuracy, we should collect large volumes of eye gaze data and refine current eye-tracking model used in our system. Lastly, the stream data provided by web-cams should be further exploited and provide analysts with high dimensional user information to comprehensively understand customers' buying behaviors. For example, by classifying users afront the shelf into different age groups while collecting their preferences over products, we could more precisely design products aim at garget groups.

4 Conclusion

As a natural customer attention tracking device, our proposed smart shelves equip marketing racks with eyetracking technology. WatchYouWatch effectively solves the issues and cognitive biases of surveys and interviews in analyzing customer preferences. While ensuring to provide a credible customer attention data over different products on the shelf, we provide users with a natural and spontaneous shopping experience. This web-cam based natural, efficient yet credible solution could have broad application on a production level in the future.

5 References

[1] A. Chowdhury, S. Karmakar, S. Reddy, S. J, S. Ghosh, and D. Chakrabarti. Visual attention analysis on mutated brand name using eye-tracking: A case study. *World Academy of Science, Engineering and Technology*, 68:1132–1135, 01 2012.

[2] J. Clement. Visual influence on in-store buying decisions: an eye-track experiment on the visual influence of packaging design. *Journal of marketing management*, 23(9-10):917–928, 2007.

[3] M. Face++. Gaze estimation. `https://www.faceplusplus.com/gaze-estimation/`. Accessed: 2018-12-17.

[4] R. Michalski and J. Grobelny. An eye tracking based examination of visual attention during pairwise comparisons of a digital product's package. In M. Antona and C. Stephanidis, editors, *Universal Access in Human-Computer Interaction. Methods, Techniques, and Best Practices*, pages 430–441, Cham, 2016. Springer International Publishing.

[5] M. Okano and M. Asakawa. Eye tracking analysis of consumer's attention to the product message of web advertisements and tv commercials. In *2017 5th International Conference on Cyber and IT Service Management (CITSM)*, pages 1–5, Aug 2017.

[6] P. M. Podsakoff, S. B. MacKenzie, J.-Y. Lee, and N. P. Podsakoff. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of applied psychology*, 88(5):879, 2003.

[7] A. Poole and L. J. Ball. Eye tracking in hci and usability research. In *Encyclopedia of human computer interaction*, pages 211–219. IGI Global, 2006.