APPLICATION OF OPENPOSE ALGORITHM TO DETECT CONSUMER BEHAVIOR IN STORE

JINGJING LI¹, JIE ZENG¹, KEYU HOU¹, JIN ZHOU^{1,2*}, RUI WANG³

¹*Sichuan University, College of Biomass Science and Engineering, National Engineering Research Center of Clean Technology in Leather Industry, Section of Chengdu No. 24 Southern Yihuan, 610065, Chengdu, China, 625368881@qq.com*

²*Science Lab, Zhejiang Red Dragonfly Footwear Co., LTD., Zhejiang Province, Wenzhou, 325100, Wenzhou, China, zj_scu@scu.edu.cn*

³*Chengdu Institute of Product Quality Inspection Co. Ltd, 610000, Chengdu, China, wangruicd@sina.com*

Due to the importance of offline consumer behavior, more and more people had begun to study consumer behavior in store. In offline consumer behavior research, the application of video analysis technology was the most direct and convenient. Recognizing human posture was a key technology in video analysis. The OpenPose algorithm was one of the advantageous technologies that could accurately recognize multi-person poses in different environments in real time, so we used it innovatively to study consumer behavior in store. We hope to develop the potential of this application in the research of consumer behavior in store in the footwear retail industry by the technical advantages of the OpenPose algorithm. In our study, we first used an OpenPose algorithm to estimate multi-person pose and detection behavior, and then processed and recognized the videos collected in the store. We collected a week's surveillance video of a Red Dragonfly offline store from July 10 to July 16, 2020 in China. The specific process was to calibrate the area in the selected camera screen, then the algorithm performs identification and detection, and finally output in-store consumption Behavioral data. Our research results not only verified the feasibility of this application in offline retailing stores, but the data results also indicated that consumers tend to enter the store from the right, staying concentrated in the middle and back of the store. These results may be affected by the store space, product display, and staff guidance and reception.

Keywords: OpenPose; Footwear; Consumer Behavior.

INTRODUCTION

Online retail had been hit hardly by offline retail, but it could not replace offline shopping (Chen, 2020), especially in footwear retailing industry (Levin *et al.*, 2003). Consumers had to go to footwear retailing stores to experience the service and fit the shoe so as to buy a suitable product (Ferreira, 2015). Consumer behavior digital could provide brands valuable information, this information could provide consumers with better service and experience. So convenient and accurate digital information collection methods had been widely studied by people (Frontoni *et al.*, 2013, Marginean *et al.*, 2019).

There were two main digital information collection methods (Mancini *et al.*, 2013), one was RFID technology, another one was video analysis technology (Chen, 2014). Marina Kholod *et al.* (Jung and Kwon, 2011) used RFID technology to collect consumer behavior in store and realize information identification of footwear products. Emanuele Frontoni *et al.* (2013) used RGBD video analysis technology to track consumer behavior in store, they successfully studied the interaction between customers and products on the shelf.

In these methods, video analysis technology (Mancini *et al.*, 2013) was a more appropriate method for digital information collection. The store monitoring camera (Liciotti *et al.*, 2014) was a common devices in store and it could be directly applied in

<https://doi.org/10.24264/icams-2020.III.11>

body pose estimation (Frontoni *et al.*, 2013), and recognizing video, predicting human posture was an extremely important part of it. The application of artificial intelligence realized the development from identifying a single person to multiple in different environments and track (Cao *et al.*, 2018). There were mainly two mainstream multi-person recognition models, one was to detect multiple people first and then estimated each person's pose, which had high accuracy but low processing efficiency (Li *et al.*, 2018), such as AlphaPose; another one was to detect the joint points first, and then determined which person each joint belongs to, such as OpenPose (Cao *et al.*, 2018). OpenPose could capture and recognize human body postures and movements in real time, it has advantages in in-store consumer behavior research. Because it was mainly used in the recognition of human skeleton points and gesture recognition (Osokin, 2018), including fitness (Noori *et al.*, 2019), action collection (Li *et al.*, 2018), etc. These above studies still had shortcomings, there was no actual application of these algorithms to the footwear and garment retail industry, and they also lack planning and long-term applications.

Therefore, the purpose of this research was first to establish in-store consumer behavior detection algorithm based on OpenPose to verify the feasibility of this new method, and then processed the data after identifying the video of the store to get the data result of in-store consumer behavior.

METHOD

Target

We used OpenPose algorithm devised an application to process surveillance video of a shoe store, which was a store of Red Dragonfly in Chengdu. Red Dragonfly was as well as a well-known shoe company in China. And we collected videos of this store from July 10th to July 16th, the daily video time was 8 am to 10 pm.

Technology Architecture

Algorithm Technology

The openpose algorithm in this research was responsible for recognizing the video datas taken by the camera and determining the consumer behavior in the store. The application process of the algorithm was to first use the 2-way convolutional neural network VGG-19-branch, then predicted the 2D vector of the key bone connection mode, and then used greedy reasoning to analyze the human body pose from the confidence map. Finally it would output the coordinates of each individual and each joint independent point. These data were the result of this research

In order to ensure the reliability of data identification, we carried out a preliminary actual verification of the accuracy of the method first. The video clipped in the same period were selected for three times of identification to obtain the results respectively, then the completeness and usability of the data were checked by comparing the manual counting results. If the video recognition result was unstable, judged the video data this time as invalid and retest until the result was stable.

Processing the Undetected Frames with Median Filter

Due to the occlusion of human joints and the defects of the OpenPose algorithm, some joints may not be detected in certain frames, and median filter was performed on the joints coordinates of these frames which appeared as both X and Y coordinates are 0. Set $PosX(k)$ and $PosY(k)$ represent the x and y coordinates at k th time respectively which were not be detected. Median filtering process was as follows:

$$PosX(k) = \frac{PosX(k-1) + PosX(k+n)}{2} (k \neq first \text{ and } k \neq end) \quad (1)$$

$$PosY(k) = \frac{PosY(k-1) + PosY(k+n)}{2} (k \neq first \text{ and } k \neq end) \quad (2)$$

Among them, $PosX(k+n)$ and $PosY(k+n)$ were the first points whose X and Y coordinates were not 0 after time k .

Digital FIR Low-Pass Filter Designed

In order to remove high-frequency jitter, it was necessary to perform low-pass filtering on the obtained continuous coordinate data. People walking frequency did not exceed 2-3 Hz, so set the passband frequency to $0 \text{ Hz} \leq f_p < 6 \text{ Hz}$. The 17-order FIR filter designed by MATLAB R2019b was used to low-pass filter the coordinates of each joint.

Data Processing

Consumer Action Judgment

The criteria for determining consumer behavior in this study were: 1) When someone entered the entrance area, it was marked as in-store passenger flow; 2) When someone stood in the consumer's attention area for more than 5 seconds, it was marked as having an intention to shoe in that area; 3) When someone stood in the consumer's attention area and raised his hand to take the shoes, it was marked as having an intention to shoe in the area; the algorithm counted the number of focus people in each area and facilitated further analysis.

Regarding the calibration method of the sales area in the camera equipment image. There were three areas in the screen: entrance, product display and consumer attention. The consumer's attention area was manually calibrated, it was a frame with the same width as the product showcase and a length of 80cm. In this study, the 5 areas in front of 5 stations were calibrated, as shown in the following figure:

Figure 1. Schematic of the area calibration of the monitoring screen in the study

<https://doi.org/10.24264/icams-2020.III.11>

Passenger Flow-Attention on Conversion Rate Calculation

Divided the number of video attentions into five categories, and accumulated the effective attention in each area according to the period, so as to obtain the number of consumer attention A of the products in each area in stores. Determined the number of attention A and the number of passengers P in the corresponding period according to the following formula:

$$A = \sum_{i=1}^1 ai \quad (3)$$

$$P = \sum_{i=1}^1 pi \quad (4)$$

Among them, the formula represents the number of follow-ups within the set time period I , the number of follow-ups per day was f_i , the number of passenger flows was p_i , i represented the serial number of each day, $i=1,2,3,\dots$, and the corresponding date serial number. From the number of follow-up times A and the number of in-store visitors P in the period, then we could obtain passenger flow-following conversion rate T . Determined the attention conversion rate T during the period according to the following formula:

$$T = \frac{A}{P} \% \quad (5)$$

RESULTS

Figure 2. The OpenPose human body gesture recognition screen in the study

We applied this algorithm in the experiment to get the recognition result graph shown in Figure 2. This method could better detect and display all the key points of the human posture in each area of the video, and successfully output the data results. According to the data in the collection period, we obtained the results in Table 1.

Table 1. In-store consumer behavior data analysis results

Category	Passenger Flow			Attention	
Daily average	1765			1302	
Passenger flow-attention on conversion rate	73.76%				
Aera	A	B	C	D	E
Passenger flow-attention on conversion rate (Daily average)	2.2%	20.7%	41.2%	8.4%	1.2%

First, the overall customer flow conversion rate could be optimized. Secondly, in the conversion rate of each area in stores, area C was the highest at 41.2%, and area A and E were extremely low. In terms of the overall distribution of consumers entering the

store, consumers preferred to enter the store on their own right. After entering the store, their trajectories were concentrated on the right side of the store, gathered in area B and C; Compared to area A and E, they were more concentrated in the middle and rear of the store, such as area B, C, and D.

DISCUSSION

In this research, we innovatively applied the OpenPose algorithm in offline retail to identify consumer behavior in store. Experiments were conducted in a real environment, it also objectively verified the reliability of the application in the store. After processing the surveillance video datas, we finally successfully obtained the data result of in-store consumer behavior.

From the application's successful verification, compared with RFID and other video analysis technology to study consumer behavior in store, the advantages of the technology used in our study were high efficiency, high accuracy, cost reduction, and reduction of human and material resources. Zhe Cao *et al.* (2018) also verified that.

The in-store consumer behavior data result showed the trajectory trend of consumers entering the store to the right, middle and back. It was verified the finding of Groeppel-Klein & Bartmann (2007), they also suggested that customers were used to entering the store from the right and then starting to search for products. The reason for this was related to the spatial layout of the store. In one hand, customers would avoid the not-wide locations and the location of the cashier counter (Ebster, 2011); On the other hand, it may be related to the display of men's and women's shoes and shop assistant (Alexander *et al.*, 2002). Because who entered the store were mainly women, the sales staff could lead female customers to area B and C directly, as we could find from the picture, the styles of sandals in these areas were more suitable for the current season. The data results were also convenient for the store staff to judge the location of more attention based on this situation, and summarized the hot products and the display area in the store, so they could update the layout of the store in time (Frontoni *et al.*, 2013).

However, there were still some problems in this study: First, when the human body was overlapped or occluded, the single camera device collection may get less accurate video results. Secondly, the accuracy of algorithm recognition could be enhanced.

Therefore, in the future, we can increase the number of cameras in store to improve the accuracy of video results. Secondly, we can improve the speed of algorithm recognition and response, associate relevant product information when recognizing behavior, and add analysis of regional matching products to the results. It can also combine the observation and analysis of facial recognition technology to obtain more detailed information about the crowd that produces a specific behavior. Overall, the technology has a wide range of potential for consumer behavior research in the retail industry.

CONCLUSION

We proved this algorithm was a convenient and accurate method that could be applied to offline retailing stores for a long time and on a large scale. It could give the store more accurate consumer behavior data analysis so it can be more widely used in store consumer behavior research.

REFERENCES

- Alexander, G.M., Packard, M.G. and Peterson, B.S. (2002), "Sex and spatial position effects on object location memory following intentional learning of object identities", *Neuropsychologia*, 40(8), 1516-1522, [https://doi.org/10.1016/S0028-3932\(01\)00215-9](https://doi.org/10.1016/S0028-3932(01)00215-9).
- Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E. and Sheikh, Y. (2018), "OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields", Paper presented at the *IEEE Conference on Computer Vision and Pattern Recognition*, 7291-7299, <https://doi.org/10.1109/CVPR.2017.143>.
- Chen, C.-C. (2014), "RFID-based intelligent shopping environment: a comprehensive evaluation framework with neural computing approach", *Neural Computing and Applications*, 25(7-8), <https://doi.org/10.1007/s00521-014-1652-7>.
- Chen, Y. (2020), "How the Internet Affects the Development of Retail Industry", Paper presented at the *International Conference on Applications and Techniques in Cyber Security and Intelligence*, https://doi.org/10.1007/978-3-030-53980-1_141.
- Ebster, C. (2011), *Store design and visual merchandising: creating store space that encourages buying*, Business Expert Press, NY, <https://doi.org/10.4128/9781606490952>.
- Ferreira, J.P.S. (2015), "Online shopping behavior in offline retail stores: strategic value for companies?" *Universidade Católica Portuguesa*, 11(4).
- Frontoni, E., Raspa, P., Mancini, A., Zingaretti, P. and Placidi, V. (2013), "Customers' activity recognition in intelligent retail environments", Paper presented at the *International Conference on Image Analysis and Processing*, https://doi.org/10.1007/978-3-642-41190-8_55.
- Jung, I.-C. and Kwon, Y.S. (2011), "Grocery customer behavior analysis using RFID-based shopping paths data", *World Academy of Science, Engineering and Technology*, 59, <https://doi.org/10.5281/zenodo.1081169>.
- Levin, A.M., Levin, I.R. and Heath, C.E. (2003), "Product category dependent consumer preferences for online and offline shopping features and their influence on multi-channel retail alliances", *J. Electron. Commerce Res*, 4(3), 85-93.
- Li, B., Zou, J., Wang, L., Li, X., Li, Y., Lei, R. and Sun, S. (2018), "The Overview of Multi-person Pose Estimation Method", Paper presented at the *International Conference on Signal and Information Processing, Networking and Computers*, https://doi.org/10.1007/978-981-13-7123-3_70.
- Liciotti, D., Zingaretti, P. and Placidi, V. (2014), "An automatic analysis of shoppers behaviour using a distributed rgb-d cameras system", Paper presented at the *2014 IEEE/ASME 10th International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, <https://doi.org/10.1109/MESA.2014.6935617>.
- Mancini, A., Frontoni, E., Zingaretti, P. and Placidi, V. (2013), "Smart vision system for shelf analysis in intelligent retail environments", Paper presented at the *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, <https://doi.org/10.1115/DETC2013-12317>.
- Marginean, A., Brehar, R. and Negru, M. (2019), "Understanding pedestrian behaviour with pose estimation and recurrent networks", Paper presented at the *2019 6th International Symposium on Electrical and Electronics Engineering (ISEEE)*, <https://doi.org/10.1109/ISEEE48094.2019.9136126>.
- Noori, F.M., Wallace, B., Uddin, M.Z. and Torresen, J. (2019), "A robust human activity recognition approach using openpose, motion features, and deep recurrent neural network", Paper presented at the *Scandinavian Conference on Image Analysis*, https://doi.org/10.1007/978-3-030-20205-7_25.
- Osokin, D. (2018), "Real-time 2D multi-person pose estimation on CPU: Lightweight OpenPose", arXiv preprint arXiv:1811.12004, Cornell University, NY, <https://doi.org/10.5220/0007555407440748>.