

Pervasive System for Consumer Behaviour Analysis in Retail Environments

Daniele Liciotti^(✉), Emanuele Frontoni, Adriano Mancini,
and Primo Zingaretti

Dipartimento di Ingegneria dell'Informazione, Università Politecnica delle Marche,
Via Brecce Bianche, 60131 Ancona, Italy
{d.liciotti,e.frontoni,a.mancini,p.zingaretti}@univpm.it

Abstract. In the context of retailing, the monitoring of consumer behaviours is particularly important for supporting vendors in their management and marketing decisions. Many studies have been carried out about various aspects and consequences of different behaviours. However, only recently the potential of computing systems is being used for automated data collection and processing. In this work, we present a novel pervasive system able to automatically monitor consumer behaviour in front of shelves in an intelligent retail environment (IRE). Data collected are stored into a cloud server for data analysis and insights, ready to be used by a Decision Support System (DSS).

The completely autonomous and low cost system proposed in this paper is based on a software infrastructure connected to a video sensor network. A set of computer vision algorithms, embedded in the distributed RGB-D cameras, provides information concerning customer behaviour, in particular, user-shelf interactions described with temporal and spatial features. This large number of analytics allows insight deductions. The use of distributed vision sensors inside a retail environment is novel and produces really valuable data for brands and retailers.

The feasibility and the effectiveness of the proposed architecture and approach have been tested on real retail environments.

Keywords: Ambient intelligence · Computer vision · Embedded sensors · Pervasive systems · Consumer behaviour · Retail

1 Introduction

In literature, there are several researches that study the behaviour of consumers in retail environments, for example, [7, 15] and references therein. In particular, Puccinelli et al. [15] identified seven topic areas of consumer behavior research in retail environments: (1) goals, schema and information processing, (2) memory, (3) involvement, (4) attitudes, (5) affect, (6) atmospherics and (7) consumer attributions and choices. For each topic, they highlighted the most important issues necessary to be further investigated.

A common characteristic of all these studies is to do not use automated approaches for data acquisition and information retrieval. They mainly focus on consumer research and retailing from the social, psychological and marketing point of views. On the contrary, the potential of computing to improve all aspects of retail is firstly studied in deep in [9]. In particular, computer vision systems appear very useful in retail environments (as well as in other application fields), mainly for the huge amount of data and the possibility of an automatic data collection. Obviously, an increasing number of these applications [13,19] have been and are possible thanks to the strengthening of information systems, the development of more stable and efficient vision algorithms and also the higher speed and the lower price of current hardware.

Focusing on computer vision approaches of consumer attributions and choices, Chandon et al. [2] and Strandvall [16] both used eye tracking methods for measuring the value of point-of-purchase. Määttä et al. [12] classified shopper motion into four behaviour classes, distinguishing if their movement is neutral or repetitive. Another approach for video-based extraction of customer movements at the point of sale is described in [8]: their human behaviour analysis is based on the measurement of the customer trajectories inside the store and on the time spent by each person in each zone of the store. According to [14], newer video surveillance applications, not necessarily related to security issues, were developed for shopping, not only to identify anomalous activities, but also to identify people and to analyse consumer behaviour.

At the same time, other pervasive computing approaches were adopted to solve problems in retail environments. Another system developed for the retail store is “SmartStore” [6], which analyses the customer interest immediately, gathers the sensing data from large-scale area and attaches massive tiny sensors to shopping items.

Our research focuses on the implementation of a software infrastructure coupled with a hardware technology to build a pervasive computing intelligent system for detecting and analysing the human behaviour in real retail stores.

By means of video cameras and computer vision algorithms, our pervasive system detects human motion and then describes human behaviour by quantitative parameters. More in detail, the main objective is to analyse the interactions between customers and products on the shelves. Therefore, our system detects and monitors people when they are in front of a shelf, using a distributed video sensor network. This allows us to better detail the activities of consumers when they stop in a zone of the store, e.g., the objects of the shelf that are touched by each person.

The installation of the system in several parts of the store provides large volumes of multidimensional data on which to perform statistics and deduce insights. The analysis of these data offers a unique possibility to better understand several crucial aspects of a retail ambient, e.g., the appealing of a product, a good positioning of different products on a shelf, the human traffic in front of each shelf.

Just because the sensor installation should be repeated in several zones of the shop to collect a significant large amount of data, we developed a system that can be easily scaled, from a single shelf installation to a large widespread grid of sensors.

Our system is able to detect all the objects on the shelf that interact with the customer using only one RGB-D sensor for each shelf, while in [7] the activity recognition needs an RFID sensor mounted on each object to be performed.

We developed a software infrastructure that is able to automatically detect, measure and store crucial information for a retail ambient. In particular, our pervasive system does not need to interact with customers to retrieve the desired information, as for the cases of the interactive display [17] or the mobile phone [5], where direct customer interactions with the systems are used to exchange information between shoppers and the retail ambient. In our case, information are collected automatically through the computer vision algorithms that we developed.

Summarizing, the paper presents a smart and low-cost embedded sensor network for intelligent retail environments (IREs) able to identify customers and to analyse their behaviour and shelf interactions. Major characteristics of this system are the general and easily scalable architecture really focused on the retail environment application and the very precise and reliable computer vision algorithms, which are able to run efficiently in low cost hardware and to collect automatically several relevant information.

The paper is organized as follows: in Sect. 2 we describe the IRE architecture and application requirements; Sect. 3 describes the implemented computer vision algorithms, focusing on people monitoring and detection of user-shelf interactions; experimental setup and results are presented in Sects. 4 and 5, respectively; finally, conclusions and future works are presented in Sect. 6.

2 IRE Architecture and Application Requirements

The hierarchical architecture of the proposed pervasive retail environment is shown in Fig. 1 along with the information provided at each abstraction layer. Sensor nodes, able to measure autonomously a part of the environment, are logically connected to the concept of shelf, multiple shelves are part of a store and, finally, several stores are part of a retailer chain. The general idea is based on several aggregation layers that provide to the system different information, from raw data to high level data analytics and insights.

At the single camera node level only raw data are available; a first data processing to provide interaction maps occurs at the shelf level. Multi camera analysis is also functional to perform flow comparisons in different areas of the store. At the top level general insights, store comparisons, store optimizations and re-design can be performed by retailers.

The functional requirements of the system, concerning what the system is able to do, its expected behaviour and which are its input/output functions, are:

1. counting the number of people “passing by” the camera vision area;

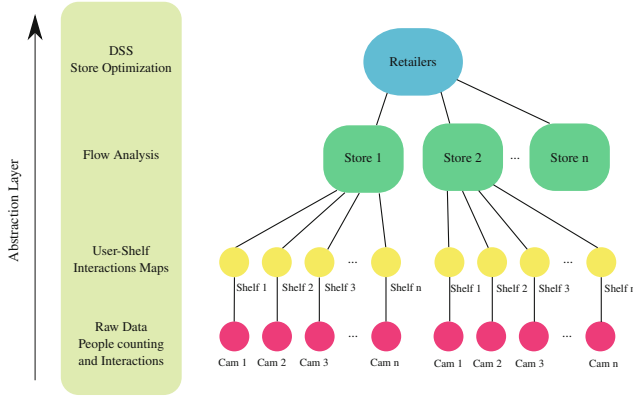


Fig. 1. The hierarchical architecture and the information provided at each abstraction layer of an intelligent retail environment.

2. for each person detected by the camera, storing the input/output position and the permanence time in the camera vision area;
3. storing all the interactions with the shelf:
 - (a) who performed the action;
 - (b) which product was touched;
4. sending and saving data in a remote database (cloud) available for statistical analysis;
5. from any remote location, controlling RGB-D camera parameters:
 - (a) video streaming visualization;
 - (b) redefinition of shelf area;
 - (c) software upgrade;
6. restoring data from a backup.

Non-functional requirements establish constraints on the system or on its development process, defining how it should be achieved. For the implementation of the system, precise constraints have been imposed:

1. vision sensors should be ceiling mounted, which limits both their size and the total weight of all components to be supported by the structure;
2. cameras should be connected to a main router via a wireless network and accessible from outside of the store for remote control;
3. the whole system should be low cost and scalable.

3 Computer Vision Algorithms

The whole system can be seen as a big sensor network where each node is a micro system that analyses the consumer behaviour in front of a specific shelf of a store. Each node consists of an embedded system that includes an RGB-D sensor, ceiling mounted and looking to the scene from the top, and a software component

to send the calculated information to the cloud. Adding a new node/shelf does not require structural modifications of the entire system, so that it is possible to install several RGB-D sensors in every store.

In this section we focus the attention on the two vision algorithms that we have already developed, processed by every node: people monitoring and user-shelf interactions.

To build a system able to detect automatically all the relevant data for a retail environments other algorithms are being developed. In particular, our current researches concern also person re-identification, in order to create an architecture comprehensive of highly-integrated systems able to totally monitor consumer behaviour in a retail environment. The first problem, even if highly studied in the literature [1], is not as much treated for top-view images, where the re-identification of each visitor is made using features like height, colors of hairs and clothes of people that are in the camera field. The second research, on the base of our previous work on Ambient Assisted Living field [11], intends to propose an audio framework able to recognize voice commands by continuously monitoring the acoustic retail environment. Echo and interference cancellation algorithms guarantee a reliable keyword extraction, which is a very important aspect from the store point of view, since it allows to know and follow customer preferences.

3.1 People Monitoring

In this part we describe our algorithm for people monitoring, which consists of two steps: background subtraction and people detection.

As regard background subtraction, in this work we implemented the algorithm of Zivkovic [20], because it ensures a lower computation time than the method we used in [10, 18].

For people detection we used the segmentation algorithm named Multi-Level Segmentation, which has in input the foreground depth image (the output of the background subtraction algorithm) and in output the highest point of each person. The algorithm is explained in detail in the pseudo-code Algorithm 1. The `MULTILEVELSEGM` function has in input the foreground image ($f(x, y)$). First of all, `FINDPOINTMAX` function calculates the highest point of whole image (max) and its coordinates ($point_{max}$). In line 3, the *level* counter assumes the *threshold* value, that is a fixed value corresponding to average height of a human head (we adopted the value 10 cm). So, the number of segmentations is strictly related to the height of the tallest person. The output condition of the while loop is when the segmentation level becomes negative (above the floor). The `SEGM` function yields in output a binary image with blobs representative of moving objects that are above the segmentation level ($max - level$). This binary image is the input of `FINDCONTOURS`, an OpenCV function that returns a vector of points for each blob. Then, the `FILTERCONTOURS` function deletes noise (blobs with a little dimension and/or a bad shape). The **for** loop from line 8 to line 14 inserts in the vector *points* the highest point/depth value (`FINDPOINTMAX` function) of each blob identified by means of the `FILTERMASK` function. Finally,

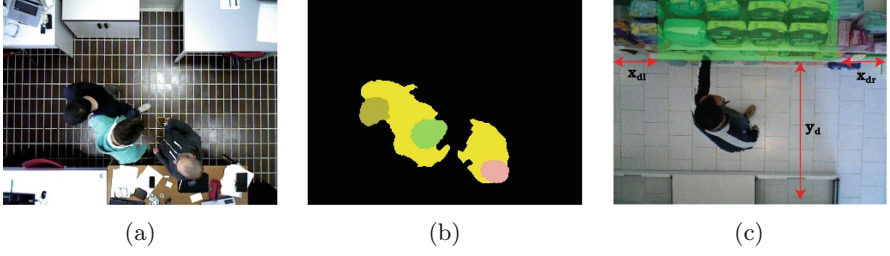


Fig. 2. Head recognition (2b): different colors of the blob highlight the head of the people detected in the scene (2a). Parameters defining a shelf zone (2c) (green area). (Color figure online)

MULTILEVELSEGM function returns a vector with all maximum local points. The length of this vector indicates the number of people that are in the image.

The Multi-Level Segmentation algorithm intends to overcome the limitations of the binary segmentation method proposed in [10] in case of collisions among people. In fact, using a single-level segmentation, in case of a collision, two people become a single blob (person), without distinguishing between head and shoulders of the person. By using this new approach, when a collision occurs, even if two people are identified with a single blob, the head of each person is anyway detected, becoming the discriminant element. Figure 2 highlights the head of each person obtained by the Multi-Level Segmentation algorithm: different colors highlight the head of a person detected by the camera. In case of collisions (Figs. 2a and b) the yellow blob contains two people and both heads are also detected.

3.2 Detection of User-Shelf Interactions

In this subsection, we focus on the algorithm for the detection of interactions with the shelf. A shelf zone is the part of the store interested by an interaction

Algorithm 1. Multi-Level Segmentation algorithm

```

1: function MULTILEVELSEGM( $f(x, y)$ )
2:   ( $max, point_{max}$ ) = FINDPOINTMAX( $f(x, y)$ )
3:    $level = threshold$ 
4:   while ( $max - level$ ) > 0 do
5:      $f_{level}(x, y) = SEGM(f(x, y), (max - level))$ 
6:      $contours = FINDCONTOURS(f_{level}(x, y))$ 
7:     FILTERCONTOURS( $contours$ )
8:     for each contour  $i \in contours$  do
9:        $f_{mask}(x, y) = FILTERMASK(f_{level}(x, y), i)$ 
10:      ( $v_{max}, p_{max}$ ) = FINDPOINTMAX( $f_{mask}(x, y)$ )
11:      if  $p_{max} \notin points$  then
12:         $points.PUSHBACK(p_{max})$ 
13:      end if
14:    end for
15:     $level = level + threshold$ 
16:  end while
17:  return  $points$ 
18: end function

```

between the hand of the shopper and a product in the shelf. As described in [4], it is defined by the user during the installation phase and it is characterized by the following tree parameters, written in a configuration file: the maximum distances of the left (x_{dl}), right (x_{dr}) and frontal (y_d) shelf sides from the image borders (see Fig. 2c). This setting is valid for most of the shelves of a store, but it is possible to define other areas of interest (e.g., in the case of a circular island when a cylinder-shaped configuration is needed).

The Multi-Level Segmentation algorithm, described in the previous paragraph, allowed us to detect the head and the body contours of each person. The contour of the head is used to track the movements of a person within the scene. The contour of entire body is used to identify interactions with the shelf. The three vertical planes built at the distances x_{dl} , x_{dr} , y_d from the image border and defining the shelf zone are used to detect interactions. When the contour of entire body intersects at least one of the three planes we establish that a contact occurs and so determine the 3D coordinates of the contact point.

The Contact Detection algorithm is explained in detail in the pseudo-code Algorithm 2. The FINDINTERACTIONS function has in input the depth image of the sensor and the contour vector. Each point of each contour is analysed (from line 2 to line 4) to find contact points with *shelfzone*. If a contact occurs the PUSHBACK method inserts in the vector *vec* the 3D contact point, where the third dimension corresponds to the depth value. Finally, the function returns vector *vec* with all the contact points.

4 Experimental Setup and Data Stored

The system setup is based on state of the art and low cost RGB-D sensors. For the development of this framework a single board computer is used since it is sufficiently small, suited to manage all functions and low consuming.

The camera sensor captures depth images with dimensions of 320×240 pixels at a rate up to 30 frames per second. The depth value is stored using 16 bits allowing a spatial resolution of few millimeters.

In addition to the RGB-D sensor, a Wi-Fi module is connected to the board to transmit the elaborated data to a cloud server by the home wireless network. Image-processing elaborations and transmission of only synthetic data are performed on the single board computer to respect privacy regulation.

Algorithm 2. Contact detection algorithm

```

1: function FINDINTERACTIONS( $d(x, y)$ , peopleVec)
2:   for each contour  $i \in \text{peopleVec}$  do
3:     for each point  $p \in i.\text{GETCONTOURS}()$  do
4:       if  $d(p.x, p.y) \in \text{shelfzone}$  then
5:         vec.PUSHBACK( $d(p.x, p.y)$ )
6:       end if
7:     end for
8:   end for
9:   return vec
10: end function

```

With this hardware setup the proposed algorithms are able to process the two video flows at 20 fps on the described low cost hardware. The cameras have been installed on panels in the suspended ceiling of the store. Each system yields as output a significant amount of data that are stored in the cloud, so that they can be successively analysed to extract useful indicators. The final tests in real stores have been realized installing RGB-D cameras for a time period of 3 months, in order to obtain statistically significant data.

A ground truth for algorithm accuracy evaluation was collected both in the laboratory and in some days of the real environment tests.

The indicators adopted to evaluate shopper behaviour and preferences can be:

- Number N_v of “passing by” people, that is people crossing camera field of view;
- Number V_z of visitors in each zone in which the image is subdivided;
- Number V_s of “passing by” people interacting with the shelf, where $V_s \subset N_v$;
- Number I_s of interactions for each person, with $I_s = I/V_s$, where I is the number of the interactions;
- Average visit time $\bar{T} = \sum_{i=1}^{N_v} \Delta t_i / N_v$, where the visit time Δt_i is the permanence of each person in the camera view;
- Number N of products touched;
- Duration of interactions $T_I = \sum_{i=1}^I \delta t_i$, where $\delta t_i = t_{i,end} - t_{i,init}$ is the difference between final and initial instant of interaction i ;
- Average interaction time $\bar{T}_I = T_I / I$.

5 Results

Table 1 summarizes the results obtained by monitoring 7 stores using a total of 15 cameras and for a working period equivalent to 45 months by a single camera. The values in the table refer to the most significant indicators introduced in Sect. 4. They reveal that the average visit time in front of the shelf is 6.21 s, while the average interaction time is 1.23 s. Moreover, the number of interactions for each person is higher than the number of products touched, corresponding to 1.45 interactions for each person.

5.1 Efficiency and Reliability of Algorithms

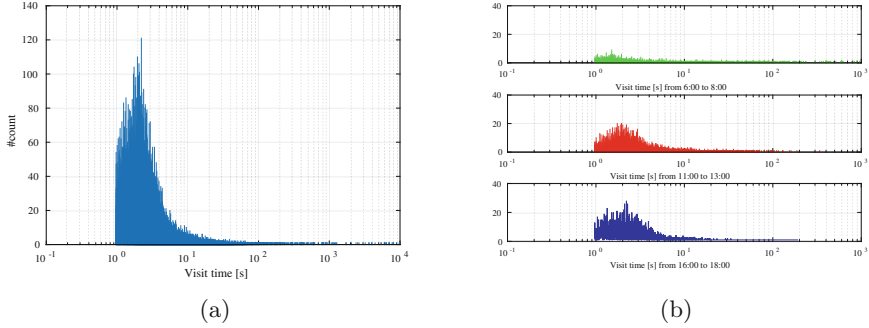
The two main requirements/capabilities that the software has to satisfy are: (i) to monitor people; (ii) to understand the occurrence of an interaction between a shopper and a shelf.

Table 1. Values of indicators for experiments in real stores.

Indicator	N_v	V_s	I_s	$\bar{T}$	N	$\bar{T}_I$
Value	87885	17762	1.45	6.21 s	25710	1.23 s

Table 2. People detection confusion and user-shelf interaction matrices.

	TP	FN	FP	Total	TPR
People detection	1110	29	11	1150	0,9902
User-shelf interaction	1050	233	254	1537	0,8052

**Fig. 3.** Visit time histogram related to the overall studied period (3a). Visit time histogram related to three different time slots: (i) 6.00 to 8.00 (green line). (ii) 11.00–13.00 (red line) (iii) 16.00–18.00 (blue line) (3b). (Color figure online)

To compare results we considered the sensitivity or true positive rate $TPR = TP/P$, where TP is the number of true positives, calculated by counting the real number of people “passing by” the camera and $P = (TP + FN)$, where FN is the number of false negatives, corresponding to the “passing by” people that the camera has not detected. The same evaluation method has been applied for establishing the correctness of the user-shelf interaction.

Table 2 shows the results of our performance analysis on 4 out of the 7 stores where the system was installed. We have checked the passages and the interactions of consumers measured by the system with the ground truth. More in detail, Table 2 corresponds to the confusion matrices of the people detection and contact detection algorithms. Since our system is built to detect only positive events (detection of people), we can not provide a value for true negative events. Also the number of negative events is unknown. For these reasons, typical confusion matrix parameters (e.g., specificity) are not listed. The sensitivity obtained was 99.02% and 80.52% for the people detection and the hand detection algorithm, respectively.

5.2 Shopper-Shelf Interaction

Once we have demonstrated the high efficiency and robustness of our algorithms, their outputs can be used for statistical analysis.

For example Fig. 3a shows the histogram of the visit time Δt_i related to the overall testing period. Each bin corresponds to an elapsed time of 1 s. The

plot shows that there are several counts (47% of the total) with a visit time smaller than 3 s, which can be easily interpreted as not-interacting people. The mean visit time higher than 3 s is equal to 6.46 s. This value, when compared to similar results related to other shelves, can describe the appealing of the shelf to shoppers. Namely, the larger is the mean value higher is the shelf attractiveness. At the same time, the total number of counts indicates if the shelf is located or not in a populated place of the shop.

Since our system detects the precise date/time when each shopper appears in the camera field of view, it is possible to make the same histogram of Fig. 3a for different time slots. Figure 3b shows an example of this for three different time slots: (i) from 6.00 to 8.00; (ii) 11.00–13.00; (iii) 16.00–18.00. During the first slot the shop is closed, hence the data refer only to shop operators. From the remaining slots we can evince that the morning time slot is more populated than the afternoon one. Such information can be very useful to better organize, for example, the staff of the shop according to the time slot with the maximum flux of buyers.

Going into the detail of the products hosted by the shelf, our system is able to storage all the interactions between the shopper and the shelf. Together with this, we can discriminate among three different types of interaction: *neutral* if the hand exceeds the threshold without taking anything; *emphpositive* when the object is picked up; *negative* when the object is put back after a pickup.

As expected the most interacted zone of the shelf is the central one, but looking at the width distribution of the interactions (Fig. 4a, top panel) it is possible to discriminate at least another peak around shelf width equal to 800 mm, that probably corresponds to an appealing product.

Figure 4 shows maps of the contact points, identified by colored zones, generated during the processing. In particular, Fig. 4a, b, c show respectively, positive, negative and neutral interactions.

Furthermore, Fig. 4 represents some example of a planogram, that is a detailed visual map that establishes the position of the products in a shelf. So to obtain the contact map, the system automatically compares the coordinates

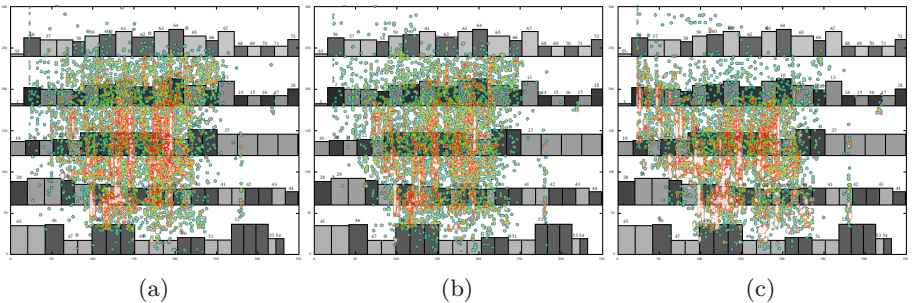


Fig. 4. Maps of interactions produced by the software in a test conducted by our research in a real environment. 2D plots showing the shelf along its width (x axis) and height (y axis) in millimetres. (Color figure online)

of contact points with ancillary information provided by the planogram management software. The planogram management software and the smart camera hardware are described in [3].

6 Conclusions and Future Works

We developed a pervasive, intelligent vision system for retail applications. It consists in a software infrastructure coupled with a low cost hardware that: *(i)* receives images from an RGB-D camera; *(ii)* elaborates the images with computer vision algorithms; *(iii)* extracts information and collects them into a database for statistical analysis and for being used by a DSS.

The implemented computer vision algorithms: *(i)* detect the people in the camera field of view; *(ii)* measure the visit time of each person; *(iii)* detect occurrences of interactions between shoppers and products on the shelf.

The system has been installed in real retail stores. The long life and real environment tests show the effectiveness of the described system and, in general, the feasibility of the proposed architecture and approach.

The efficiency of the system is defined by its capability in detecting people and shelf interactions. Results show that the people detection algorithm has a very high sensitivity (99%), while the hand detection algorithm shows a good sensitivity slightly above 80%.

The collected information can be used for several useful statistical analyses, since they enhance, e.g., the knowledge of shopper-shelf interactions and the product appeal, as shown in Subsect. 5.2.

The system can be used as part of a sensor network focused on retail reality mining, with the purpose of better understanding customer interactions in retail environments.

Our future projects are directed towards a detailed study of person re-identification using top-view images, a task necessary to assign a single and robust ID to each buyer, and audio recognition. Among several other information, this will allow us to better describe the client behaviour inside the entire shop and not only in front of a single shelf.

References

1. Bedagkar-Gala, A., Shah, S.K.: A survey of approaches and trends in person re-identification. *Image Vis. Comput.* **32**(4), 270–286 (2014)
2. Chandon, P., Hutchinson, J., Bradlow, E., Young, S.H.: Measuring the value of point-of-purchase marketing with commercial eye-tracking data. INSEAD Business School Research Paper (2007/22) (2006)
3. Frontoni, E., Mancini, A., Zingaretti, P.: Embedded vision sensor network for planogram maintenance in retail environments. *Sens. (Switz.)* **15**(9), 21114–21133 (2015)
4. Frontoni, E., Raspa, P., Mancini, A., Zingaretti, P., Placidi, V.: Customers’ activity recognition in intelligent retail environments. In: Petrosino, A., Maddalena, L., Pala, P. (eds.) *ICIAP 2013. LNCS*, vol. 8158, pp. 509–516. Springer, Heidelberg (2013). doi:[10.1007/978-3-642-41190-8_55](https://doi.org/10.1007/978-3-642-41190-8_55)

5. Güven, S., Oda, O., Podlaseck, M., Stavropoulos, H., Kolluri, S., Pingali, G.: Social mobile augmented reality for retail. In: PerCom, pp. 1–3. IEEE Computer Society (2009)
6. Iwai, M., Mori, M., Touda, H.: A marketing analysis using massive tiny sensor nodes. In: 2009 Sixth International Conference on Networked Sensing Systems (INSS), pp. 1–4. IEEE (2009)
7. Kourouthanassis, P., Roussos, G.: Developing consumer-friendly pervasive retail systems. *IEEE Pervasive Comput.* **2**(2), 32–39 (2003)
8. Krockel, J., Bodendorf, F.: Customer tracking and tracing data as a basis for service innovations at the point of sale. In: 2012 Annual SRII Global Conference (SRII), pp. 691–696. IEEE (2012)
9. Krüger, A., Schöning, J., Olivier, P.: How computing will change the face of retail. *IEEE Comput.* **44**(4), 84–87 (2011)
10. Liciotti, D., Contigiani, M., Frontoni, E., Mancini, A., Zingaretti, P., Placidi, V.: Shopper analytics: a customer activity recognition system using a distributed RGB-D camera network. In: Distant, C., Battiato, S., Cavallaro, A. (eds.) VAAM 2014. LNCS, vol. 8811, pp. 146–157. Springer, Cham (2014). doi:[10.1007/978-3-319-12811-5_11](https://doi.org/10.1007/978-3-319-12811-5_11)
11. Liciotti, D., Ferroni, G., Frontoni, E., Squartini, S., Principi, E., Bonfigli, R., Zingaretti, P., Piazza, F.: Advanced integration of multimedia assistive technologies: a prospective outlook. In: 2014 IEEE/ASME 10th International Conference on Mechatronic and Embedded Systems and Applications (MESA), pp. 1–6. IEEE (2014)
12. Määttä, T., Härmä, A., Aghajan, H., Corporaal, H.: Collaborative detection of repetitive behavior by multiple uncalibrated cameras. *Inf. Fusion* **21**, 68–81 (2015)
13. Moeshund, T.B., Granum, E.: A survey of computer vision-based human motion capture. *Comput. Vis. Image Underst.* **81**(3), 231–268 (2001)
14. Moeshund, T.B., Hilton, A., Krüger, V.: A survey of advances in vision-based human motion capture and analysis. *Comput. Vis. Image Underst.* **104**(2), 90–126 (2006)
15. Puccinelli, N.M., Goodstein, R.C., Grewal, D., Price, R., Raghubir, P., Stewart, D.: Customer experience management in retailing: understanding the buying process. *J. Retail.* **85**(1), 15–30 (2009)
16. Strandvall, T.: Eye tracking as a tool in package and shelf testing. White Paper (Toby technology) (2008)
17. Strohbach, M., Martin, M.: Toward a platform for pervasive display applications in retail environments. *IEEE Pervasive Comput.* **10**(2), 19–27 (2011)
18. Sturari, M., Liciotti, D., Pierdicca, R., Frontoni, E., Mancini, A., Contigiani, M., Zingaretti, P.: Robust and affordable retail customer profiling by vision and radio beacon sensor fusion. *Pattern Recogn. Lett.* **81**, 30–40 (2016)
19. Wang, L., Hu, W., Tan, T.: Recent developments in human motion analysis. *Pattern Recogn.* **36**(3), 585–601 (2003)
20. Zivkovic, Z.: Improved adaptive Gaussian mixture model for background subtraction. In: 2004 Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004, vol. 2, pp. 28–31. IEEE (2004)

Video Analytics. Face and Facial Expression Recognition
and Audience Measurement

Third International Workshop, VAAM 2016, and Second
International Workshop, FFER 2016, Cancun, Mexico,
December 4, 2016, Revised Selected Papers

Nasrollahi, K.; Distanti, C.; Hua, G.; Cavallaro, A.;

Moeslund, Th.B.; Battiato, S.; Ji, Q. (Eds.)

2017, X, 163 p. 86 illus., Softcover

ISBN: 978-3-319-56686-3