Measuring Customer Purchase Behaviour from CCTV Images

Measuring Customer Purchase Behaviour from CCTV Images

Dr. Andrew Newman, Dr. Daniel K. C. Yu and David P. Oulton
Department of Textiles
University of Manchester
Institute of Science and Technology
PO Box 88
Manchester M60 1QD
United Kingdom

E-mail: andrew.newman@umist.ac.uk
Daniel.k.c.yu@umist.ac.uk
David.oulton@umist.ac.uk

Abstract

This paper reports the preliminary stages of an EPSRC funded research project, which presents a new approach to analysing in-store customer behaviour with a view to optimizing space and storage performance. Using existing in-store CCTV security cameras, which were connected to time-lapsed video recording equipment, people are tracked to detect patterns of behaviour. The analysed process provides the statistical data for planning the design of merchandise placement and in-store and the arrangement of fixed and movable fittings. A clothing retailer can thus maximize sales volume by arranging, manipulating and customising the store format.

Introduction

Retail store settings are interesting places, and among other things, the varied colours, textures and general architecture of the layouts, attract consumers. Retail institutions have spent huge sums in the pursuit of improved multidimensional formats and design solutions, to attract customers into their stores. Measuring consumer behaviour is one of the best methods to understand a good deal more about the specific facets of settings and how customers interface psychologically with them. Investigating these phenomena is an enticing area of consumer research, and has been the subject of research studies [1,2,3,4]. For example, the work of Barker and Wells [5,6] pioneered research into customer behaviour drawing on the observer's interpretation of actions for the subsequent analysis. By definition this type of phenomenological approach is inherently subjective and affects all aspects of the procedure, as it attaches observer bias to recordings of behaviour even before analytical interpretation. The Ericsson and Simon [7] study in 1984 used protocol analysis to provide insight from participants' personal recollections of the process, either during or immediately following behaviour. In later work, McGrath [8] adopted an alternative research paradigm using an ethnographic approach and orientation. Most of these studies have employed either manual observational techniques, or survey and interviews that reply on consumers' retrospective recall.

More recent work in the field [9,10,11] has demonstrated that interviewer bias, the use of survey instruments and traditional methodologies in retailing have fundamentally weakens the validity of any data. In an attempt to address this shortfall, this paper proposes that unobtrusive observation through video collection substantially overcomes this problem. Recording customer behaviour using video footage from within the retail environment will provide an essential, "objective" record of movement patterns [12,13].

The capture of human movement using computer-based technology has been the focus for some development. Phillips and Bradshaw [14] can be considered as the earliest consumer researchers who used video equipment, allied to time-lapse photography to record patterns of movement within the store in 1991. Four years later, McCullagh and Thornton [15] also found that video tracking of in-store movements provides effective data concerning patterns of customer movement. The research reported in this paper, builds on prior work in an attempt to fully operationalize video capture of customer movement from retail store CCTV installations. To this end, the study advances this new methodological approach as a means of analysing customer behaviour using recognized theories of consumer behaviour.

Data Collection

After securing the services of a major clothing discount retailer known for their dynamism and attitude towards store design and development, this project recorded and analysed customer movement patterns in "real time". The commercial collaborators remit imposed restrictions on the work stressing that data collection must not affect store business. This was important consideration as the collaborator had provided the test sites and in-store access necessary to undertake the project. Data collection was automatic using previously installed ceiling mounted CCTV security cameras connected to time-lapsed video recording equipment. Different zones within the store, such as the store entrance and customer services became the areas of analysis. All cameras had been originally installed for security purposes; the initial data collection found that the focusing and angle of the in-store cameras were critical. Independent research [16,17] has demonstrated that the quality of the image generated by the video cameras was vital to successful analysis. Figure 1 shows diagrammatically the video tracking setup.

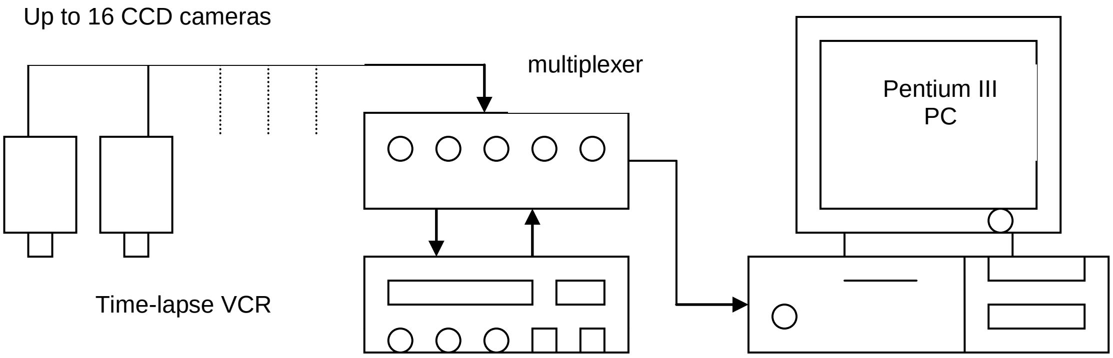

Figure 1: Diagrammatic representation of video tracking setup

Customer Tracking

Data was collected via sixteen CCTV security cameras placed strategically around a 45,000 square foot clothing store. All cameras were fixed with no panning or zooming functions. Just one of the cameras produced colour images with the remainder in black and white format. All were connected to a multiplexer, which was then linked to a time-lapse videocassette recorder. A three-hour VHS videotape was used to record twenty-four hours continuous store operation. Because of the time-lapse function, the system only provided image quality with an approximately sampling rate of 5 frames per second. This created problems for the customer-tracking objective, as the customer movement appeared jerky and therefore difficult analyze.

The raw data was digitized by using a video capture card, and stored as AVI movie format for later analysis. Data storage posed a real problem as just three minutes recording, with resolution 320 x 240 pixels in 24 bits true colour, has a file size of around 400 Mega Bytes. This presented some problems for data storage as the size of the hard disk, which must be at least 20 Giga Bytes in order to have the capacity to analyze one hour of video footage. This study used a 733 MHz Pentium III personal computer with 128 Mega Bytes RAM, and a portable zip drive for backup data. To save on hard disk space, the AVI movie formatted files were compressed into MPEG movie format. However, this process has the disadvantage of decreasing the quality of the movie, whilst taking rather long time to compress it. The authors found that it was better to retain the footage in AVI format for subsequent analysis.

Customer Counting and Motion estimation

Tracking cannot be achieved without performing motion detection. The presence of a face may be inferred by detecting differences between the current image and a reference image, by searching for large skin coloured areas or by searching for face shaped regions [18]. However, a hybrid approach, which is based on both skin colour information and shape of a face, was used in this project in order to increase the accuracy of counting. The computer identifies large skin areas from arms and legs of a customer as another customers. However, customers tend not to walk in a upright manner looking horizontally forward. In reality, the shape of a customer's head from the security camera point of view is always changing.

The traditional method for motion detection is based on frame difference, which is the subtraction of pixel values held in two different frames. However, it is only suitable for detection of continuous object movement in low speed. In the case of time-lapse video recording, continuously moving the position of the customer in each frame is relatively wide apart due the to sampling rate which is only 5 frames per second. This problem can be more easily explained with the help of simple drawings. Figure 2 below is based on normal video footage using one camera for recording. The object (in this case a figure) moves forward from frame 1 to frame 2 creating a small displacement. When pixel values from frame 1 are subtracted from frame 2, the (minimal) change is shown on screen. If the object makes a very small movement, it is possible to subtract frame 1 by another frame, for example frame 5, in order to achieve an acceptable detection result. In Figure 3, the video footage is based on time-lapse recording method with several cameras connected to a multiplexer. The same test object is shown moving forward with a much larger displacement. In this sample the difference between frames is too far apart – a factor that may result in detecting two changes rather than the incremental change required.

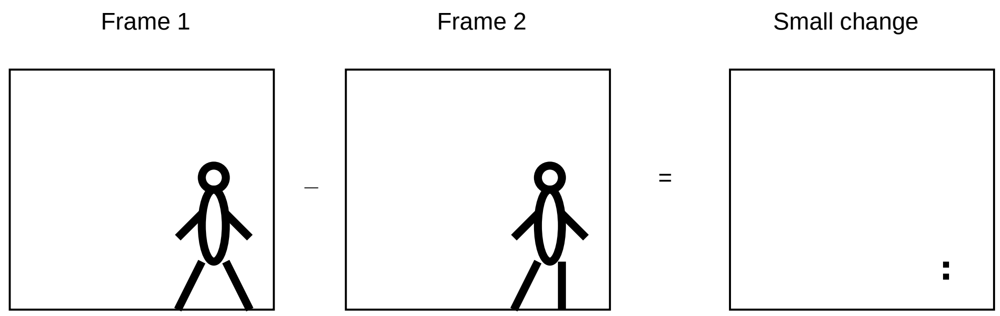

Figure 2: Motion detection based on normal video footage.

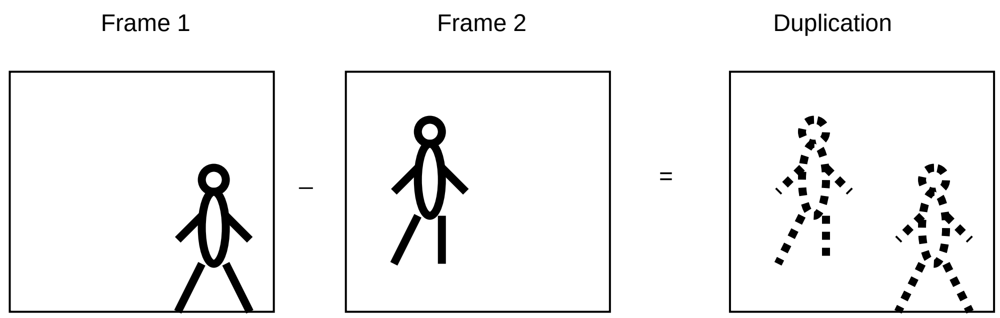

Figure 3: Motion detection based on time-lapse video footage.

As Figure 3 demonstrates, the time lapse effect created a inconsistency that necessitated the development of a new algorithm. This used the principle of motion estimation. Mattison [19] defined motion estimation as a process to determine which pixel blocks contain information representing parts of the image that have moved and how far they have moved. Thus by taking a pair of images from our video sequence and dividing each into a series of sub-images (e.g. 10x10 blocks of pixels), it was possible to determine which blocks represent parts of the picture that have moved and which have remained static. This illustrated in Figure 4 below.

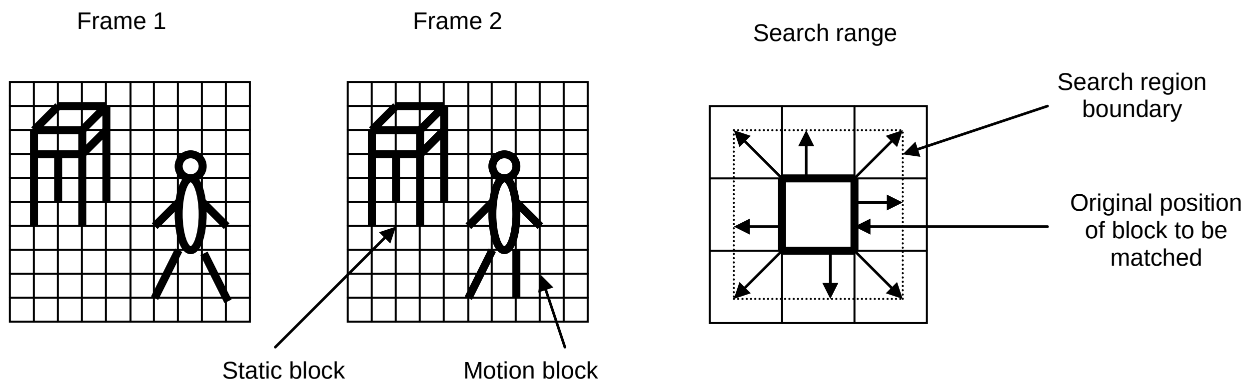

Figure 4: Conceptualization of block matching process for motion estimation.

The block matching process is conceptually quite simple although it may be fairly expensive in terms of computation time. By mapping the block from the previous image to various positions on the new image, and calculating the mathematical difference at each position, it is possible to determine mathematically whether the difference at a given location is low enough to constitute a match. It is very time consuming to search a large area of the image for each motion vector calculated. A digital mask filter was therefore developed to control the search area. The radius of the digital mask filter is a key factor in the matching process. In the case of high quality images from a normal video footage, a radius of 5 to 20 pixels is sufficient to obtain a match. However, for the analysis of customer movement from time-lapse video footage, the radius has to be increased to 200 pixels or more in order to secure a match when poor images are used.

Multi-customer tracking

The above algorithm can be used to track several customers inside the store at the same time. Two examples are provided in Figures 5 and 6 below for illustrative purposes. Figure 5 is based on time-lapse video footage and Figure 6 on normal video footage. For identification purposes, each entity in the video footage is given a numerical value to represent the point from which they were first tracked. We can see how the second example (Figure 6), which is based on normal video footage using a high quality camera, gives a more complete picture of customer behaviour than Figure 5.

Figure 5: Multi-customer tracking based on time-lapse recorded video footage

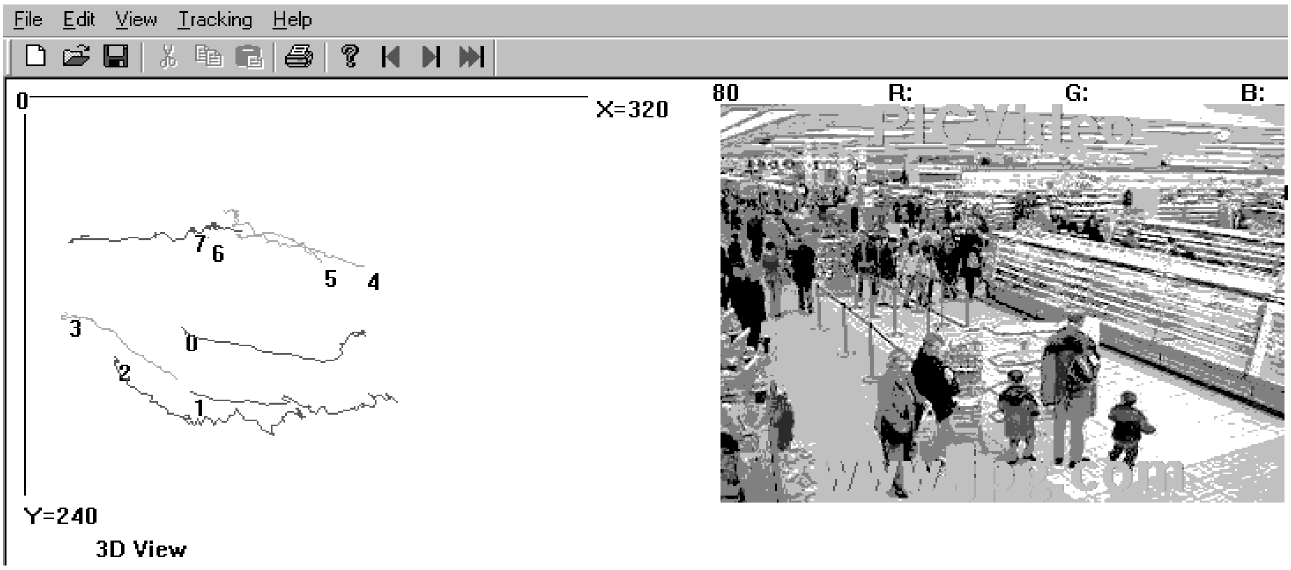

Figure 6: Multi-customer tracking based on normal recorded video footage.

In Figure 5, staff ‘0’ and staff ‘1’ are seen walking in front of the customer service counter. Customer ‘2’ walks from the entrance area into the store. Figure 6 shows eight customers being tracked at the same time using a high quality camera. Track ‘0’ depicts a male subject walking in a relatively straight line and subsequently turning to observe something of interest on the supermarket shelving. At this point the subject stopped to browse; a factor that is represented by dwell period. Entities “1” and “2”, are two young children which explains the wavy patterns of movement, as children usually display excitement and erratic behaviour during the shopping experience.

Customer arrival patterns

Other uses of this research focus on the categorical data that may be routinely collected during store opening hours. Figures 9 and 10 provide an illustration of the functional side of counting customers in-store. We can see from the data that customer arrival patterns tend to cluster round the middle of the day with low footfall at start and close of business. These may provide patterns of movement that help to schedule staffing and merchandise replenishment processes.

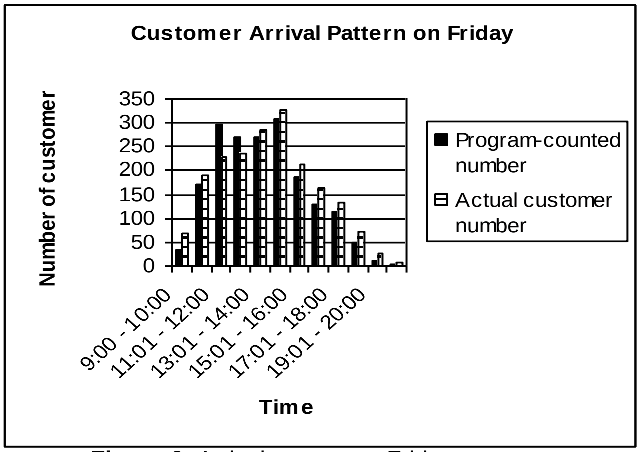

Figure 9: Arrival pattern on Friday.

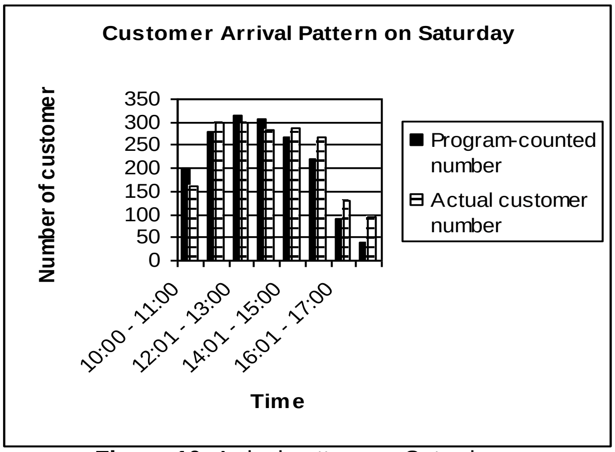

Figure 10: Arrival pattern on Saturday.

Conclusions

This paper has described the data collection mechanism and two methods of processing the data in order to measure customer purchasing behaviour. With suitable camera positions, the motion estimation algorithm can track customers from the moment of their arrival at the entrance to the last minute they are leaving the store. It is very important to know customer traffic flow inside a retail store, as the traffic pattern can help to analyze how the store's layout and general format can influence consumers' buying behaviour. From research in retailing, marketing and environmental psychology we know that a store's atmosphere is likely to positively or negatively influence a shopper's emotional state. This in turn can lead to increased or decreased shopping activity. For example, consumers' moods which are made up of the three emotional states: pleasure, arousal and dominance, have been shown to be influential on the way in which consumer spends time in the store [2]. However, this and other retailing research has failed to provide retailers with practical solutions to the problem of enhancing the shopping experience. Practitioners expect solutions to problems of hot and cold spot and require new knowledge to support the creation of store interiors. With an accurate knowledge of layouts and fixtures, and a theoretical understanding of the impact of such factors, it will be possible to record and predict the influence of design features like fixtures, entrances and walkways on customer behaviour. For example, this study demonstrated that subjects tend to dwell close to points of interest and this is similar to observed behaviour noted in the biological sciences [20].

This research therefore embraces a new and innovative methodological approach, which will use computer simulations to experiment with and reengineer store layouts. The intention is to lessen shopping stress induced by ambiguous settings. An example of this is when a consumer is faced with frequent or unexplained (in terms of signage) layout changes, and/or changes in merchandise position so creating major disincentives to shop. Clear and legible store designs, incorporating aisle design fixtures and merchandise, will reduce the uncertainty that consumers associate with shopping in large floor spaces. Customers can focus on the purchase rather than develop their navigation skills.

Using CCTV footage, the customer counting algorithm can detect number of customers anywhere in the store, and at any period of time. As Figures 9 and 10 demonstrate, the algorithm can achieve a 94.5 percent accuracy when counting customers in-store. Using this categorical customer movement data, simulations will establish the hot and cold spots in the sample retail store(s). The retail partner can then strategically arrange store architecture according to the behaviour their management wishes to evoke (e.g. circulation, left or right movement, accessibility). This may center on greater access to merchandise that would ultimately lead to increased conversion (sales).

A retailer's overall goal is to maximize profits through increased sales by way of cost-effective store design. The two algorithms demonstrated in this paper can be used as part of new methodological approach that will help to achieve this goal. The next stage of this project is to systematically record customer movement on video footage, and classify store layouts in line with these data.

References

1. Donovan, Robert J and Rossiter, John R 'Store atmosphere: An environmental psychology approach'. *Journal of Retailing*, 58, (1982) pp. 34-57.

2. Donovan, Robert J, Rossiter, John R, Marcoolyn, Gilian, and Nesdale, Andrew 'Store atmosphere and purchasing behaviour'. *Journal of Retailing*, 70, 3, (1994) pp. 283-294.

3. Baker, Julie, and Parasuraman, Dhruv Grewal A. 'The influence of store environment on quality inferences and store image'. *Journal of the Academy of Marketing Science*, 22, 4, (1994) pp. 328-339.

4. Zimmer, Mary R and Golden, Linda L, 'Impressions of retail stores: A content analysis of consumer images'. Journal of Retailing, 64 (fall), (1988) pp. 265-293.

5. Barker, R G 'Explorations in ecological psychology' American Psychologist, Vol.20 (1965) pp.1-14

6. Wells, W D and Lo Scutio, L A 'Direct observation of purchasing behaviour' Journal of Marketing Research, Vol.3 (1966) pp. 227-33

7. Ericsson, K A and Simon, H A 'Protocol analysis: verbal reports as data' MIT Press, Cambridge, MA

8. McGrath, M A 'An ethnography of a gift store: trappings, wrappings, and rapture' Journal of Retailing. Vol. 65 No. 4 (1989) pp. 421-9

9. Newman, Andrew J, Davies, Barry J and Dixon, Gordon, 'A theoretical model approach to terminal shopping syndrome: A future challenge for UK airport operators?' 24th. European Marketing Academy Conference, ESSEC, (1995) Paris (May).

10. Newman, Andrew J, Davies, Barry J and Dixon, Gordon 'Identifying the social and physical environs affecting persons in UK airports: An archetypal model of consumption advanced' Marketing Education Group Conference, University of Bradford (July 1995).

11. Newman, Andrew J, 'Consumption and the airport setting: 'Find' the time to go shopping'. 5th. International Conference on Recent Advances in Retailing and Services Science, Baveno, Italy (August 1998).

12. Newman, Andrew J 'Devising a new methodological approach to retail format planning' 6th International Conference on Recent Advances in Retailing and Services Science, Puerto Rico (1999) July.

13. Ivor, Church and Newman, Andrew J, 'Simulating Production-line Service Operations: A Systematic Approach to Process Management'. The British Food Journal, 102 (5/6) (2000) pp. 398-405.

14. Phillips, H and Bradshaw, R ' Camera tracking' Management Research News, Vol.14 No.4/5 (1991).

15. McCullagh, M J and Thornton, S 'In-store surveys of customer movement and behaviour' University of Nottingham (1995).

16. Oulton, D P and Yu, K C Daniel 'Digital camera calibration validation report' MAFF Project AFM/65 (1999).

17. Yu, K C Daniel and Oulton, D P 'Colour difference acceptability for calibrated digital images' The Imaging Science Journal (date to be advised).

18. Mattison, Phillip E 'Practical digital video with programming examples in C' John Wiley & Sons, Inc. (1994).

19. Morris, Tim 'Multimedia systems: delivering, generating and interacting with multimedia' Springer (2000).

20. Wratten, Stephen D. (1994), 'Video techniques in animal ecology and behaviour' Chapman & Hall, London (1994).