Audience Monitor – an Open Source Tool for Tracking Audience Mobility in front of Pervasive Displays

Ivan Elhart[1], Mateusz Mikusz[2], Cristian Gomez Mora[1], Marc Langheinrich[1], and Nigel Davies[2]
[1]Faculty of Informatics, Universitá della Svizzera italiana (USI), Lugano, CH
[2]School of Computing & Communications, Lancaster University, Lancaster, UK
{ivan.elhart, cristian.gomez.mora, marc.langheinrich}@usi.ch {m.mikusz, n.a.davies}@lancaster.ac.uk

ABSTRACT
Understanding an audience's behavior is an important aspect of evaluating display installations. In particular, it is important to understand how people move around in the vicinity of displays, including viewer transitions from noticing a display, through approach, to final use of the display. Despite the importance of measuring viewer mobility patterns, there are still relatively few low-cost tools that can be used with research display deployments to capture detailed spatial and temporal behavior of an audience. In this paper, we present an approach to audience monitoring that uses an off-the-shelf depth sensor and open source computer vision algorithms to monitor the space in front of a digital display, tracking presence and movements of both passers-by and display users. We believe that our approach can help display researchers evaluate their public display deployments and improve the level of quantitative data underpinning our field.

Author Keywords
Public Displays; Display Deployments; Display Evaluation; Audience Monitoring.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (HCI): Miscellaneous.

INTRODUCTION
The number of public displays in our environment has been increasing in recent years. These display deployments have been created to achieve different tasks and support diverse research questions and objectives. However, evaluating public display installations is no trivial task [2]. A common way to evaluate a new installation is through understanding its audience behavior, e.g., how people move through the space, when and how passers-by notice the display, how potential users approach it, and how users interact with the display [4][15][26]. The spatial and temporal aspects of audience behavior can be complex to capture. Many deployments thus opt for manual observations and ethnographic studies over short periods of time, e.g., two evenings on a single weekend [2].

PerDis '17, June 07-09, 2017, Lugano, Switzerland

ACM ISBN 978-1-4503-5045-7/17/06…$15.00
http://dx.doi.org/10.1145/3078810.3078823

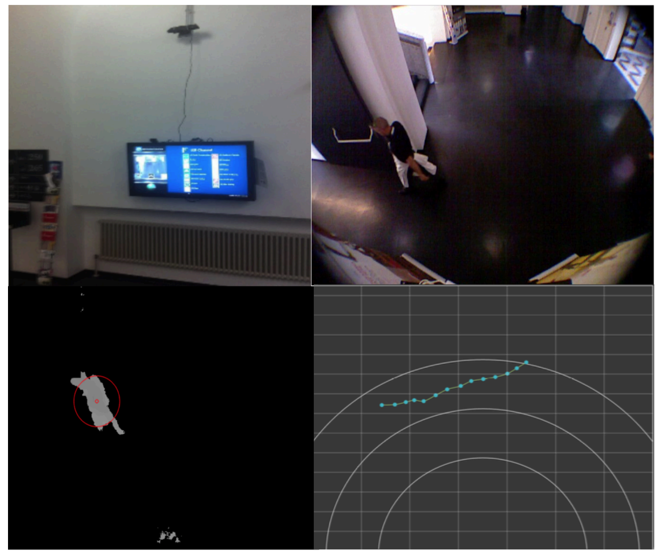

Figure 1 – Audience Monitoring Tool for Public Displays. Top left: Kinect sensor mounted on the wall above the existing public display. Top right: View of the scene from the sensor (video was recorded only during the evaluation). Bottom left: Depth view of the scene and detection of passer's-by presence. Bottom right: Visualization of the passer's-by path.

Some automation of audience monitoring is possible. For example, many display systems and installations have used video analytics such as face and eyeball detection techniques to collect information about presence and overall numbers of display users [21], interactions with display content [20], or to estimate gender and age of the audience [10]. However, as indicated by Williamson and Williamson [27], there is still a lack of tools for evaluating research display installations in the wider context of audience behavior and their movements through the space. To this end, Williamson and Williamson developed a tool that utilizes a webcam and open source computer vison techniques to capture and visualize pedestrian movements around displays from a bird's eye perspective [27]. In contrast to both the narrow field-of-view characteristics of most video analytics systems for digital signage, and the large yet less accurate analysis of human mobility through physical spaces, we explore an alternative approach that aims to provide accurate presence detection in the broad vicinity of the display, while capturing no personally identifiable information.

Our main objective was to design and develop an audience monitoring tool that can capture detailed special and temporal information regarding audience behavior using anonymous depth data and open source algorithms. The use of a depth sensor enables the creation of a system that can be used at different installation positions and angles (not only on the ceiling, but also on the display) due to a possibility of rotating the three-dimensional representation of the scene to compensate for the installation angle and minimize occlusion problems.

Here, we present our experiences of the design, development, evaluation, and use of the audience monitoring tool. We believe our approach can provide display researchers with a potentially low-cost and easy to deploy solution, using a Microsoft Kinect sensor as a state-of-the-art depth camera. The source code of the tool is publicly available on GitHub (https://github.com/elhart/audienceMonitor).

RELATED WORK

Our related work lies at the intersection of audience detection and tracking, and the existing approaches for evaluating digital signage systems.

Digital Signage Analytics

The improved technologies and algorithms in video analytics also boosted digital sign analytics and allowed both researchers and commercial entities to gain detailed insights into how viewers interact with their displays—and beyond the display. Commercial products such as Intel Anonymous Video Analytics (AIM) suite [14] and Fraunhofer Avard [12] provide detailed insights into the immediate audience demographics through video processing in terms of audience counts, gender and age estimations, dwell and view times, and distance to the display. Researchers have used data gained from such video analytics tools to understand the effectiveness of their displays. For example, Farniella et al. [10] developed a framework for running targeted advertising campaigns for a currently present audience in front of a display. As an additional step, Tian et al. combined on-screen analytics data (captured by Intel AIM) with sales statistics from retail (captured directly from point-of-sales terminals) [25]. The authors were able to use the gained information to measure the "effectiveness" of the display by mapping demographics groups who purchased products with sales statistics. The learned insights were then used for improved advertising on pervasive displays.

To understand the interaction of viewers across several displays, Gillian et al. [13] designed and deployed a framework that is capable of recognizing viewers across displays (using both depth-cameras and additional sensing functionality). While this system provides analytical insights into how viewers interact with content, it also provides additional features to the user and personalized content across a range of displays.

While such techniques detect display users and can estimate users' distance to the display, they usually do not continuously track audience mobility behavior.

Audience Detection and Tracking

Audience detection and tracking have been popular topics in computer vision for crowd detection and estimation [5], as well as the detection and tracking of individuals within crowds [1]. However, in the domain of public displays, there are only few approaches for automatic audience mobility detection and tracking.

Previously, display systems and installations have used face and eyeball detection to identify the presence and count the number of display users. For example, *Ubi-hotspots* use a webcam and simple face detection to count the number of display users to augment touch interaction logs [21]. *FunSquare*, an application developed for Ubi-hotspots, used the same user count data to automatically generate new display content [18]. Also, *ReflectiveSigns* used face detection and eye tracking to estimate when users interacted with the content, which influenced automatic content scheduling [20]. In addition, *Proxemic Interaction* used a system of six infrared cameras to estimate the position of passive infrared reflective markers worn by the users (e.g., a wearable marker attached to a user's hat). The system was able to detect position, identity, movement, and orientation of the users and influence interaction and content presentation on the display [3].

Recently, Williamson and Williamson noted the lack of display evaluation tools for display researchers that can capture a wider context of audience mobility [27]. They were not only interested in detecting display users, but also in understanding how both display users and passers-by move through the space. This information helped the authors to understand how passers-by notice and approach the display, and how a new display installation can potentially disrupt the usual flow of people in public space. In order to capture audience mobility, the authors designed and developed *Pedestrian Tracker*, an evaluation tool for display installations that uses a ceiling-mounted webcam and open source computer vison techniques, such as motion detection and background subtraction [27]. The authors demonstrated the use of the tool through several in-the-wild display deployments and showed the validity of using such an approach to detect and track pedestrians with a relatively high detection rate of 68%.

A higher detection rate of passers-by can be achieved using depth information and three-dimensional representations of the scene without collecting and processing any personally identifiable information. Seer et al. used multiple Microsoft Kinect devices, mounted on the ceiling of a long corridor, and custom-built computer vision algorithms to detect and monitor the flow of passers-by [22]. Using the depth information from multiple sensors, the authors were able to track passers-by with detection rates of up to 94%.

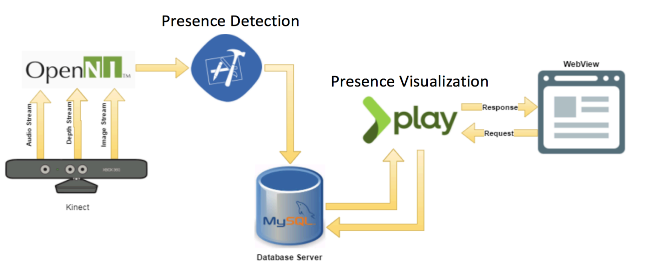

Figure 2 – Architecture Diagram of the Audience Monitoring Tool

In addition to potentially increasing the detection rate, a three-dimensional representation of a scene presents an opportunity to transform, rotate, and view the same scene from different perspectives and view angles [16][23]. This allows for sensor installations not only on the ceiling (to reduce occlusion problems), but also at different positions, heights, and viewing angles. Mounting sensors at different positions and angles can potentially support a wider range of installations, i.e., when mounting equipment on the ceiling is not possible or permitted, such as high ceilings or protected buildings.

SYSTEM DESIGN AND IMPLEMENTATION

In order to detect and monitor the audience in front of a digital display, we designed and implemented an audience monitoring tool that uses a depth sensor and a combination of open source computer vision and web visualization techniques. The tool uses an off-the-shelf Microsoft Kinect device to obtain depth information of the scene. It consists of two core modules: *presence detection* for collecting and analyzing the raw sensor readings, detecting the presence of people, and storing detection and tracking information into a database; and *presence visualization* that gives insights into the collected mobility datasets, as shown in Figure 2.

Implementation

The *presence detection* module was implemented using OpenNI (https://github.com/OpenNI), a framework for obtaining raw sensor data from the Kinect. The sensor provides 640x480 pixels of 16-bit depth information, at the rate of 30 frames per second. The module has been developed in C++ using the Xcode development environment and Mac OSX (https://developer.apple.com/xcode/). It uses the OpenCV (http://opencv.org) computer vison library for analyzing raw depth frames in real-time and detecting the presence and movement of passers-by. The detection algorithm is an extended version of a simple area detector that can filter data based on circularity, convexity, inertia, and the size of detected area. Each detected presence is stored in a database containing a unique id, timestamp, x and y coordinates of detection, and the size of the detected area. In addition, the module uses the open source GLFW openGL library (http://www.glfw.org) for transforming and rotating the three-dimensional view of the scene depending on the installation position and angle.

The *presence visualization* module takes the presence information stored in the database and provides a web-based visualization. The module has been implemented using Play, a Java based web application development framework (https://www.playframework.com/), and D3.js, a JavaScript visualization library (https://d3js.org/). The module allows the visualization of individual paths by presence id, visualization of audience presence at a specific time instance, or over a specific time interval. The module also provides histograms of the presence detection over weeks, days, and hours of the deployment, sum and average time duration of passers-by in front of the display, and distance representation using a heat map, as shown in Figure 3-7.

Installation and Deployment

We installed the tool above an existing touch-enabled public display in front of the university canteen located in a busy hall of our main university building. The Kinect device was mounted vertically on the wall approximately 1.5 meters above the display at a 30^0 angle (see Figure 1 – top left). We used a Mac Mini computer mounted behind the existing display to run all software components and store captured mobility data in a local MySQL database.

The existing display where we installed the audience monitoring tool is a part of a network of four interactive and open public displays [6] located at our university [8]. Each display has been running a set of interactive display applications that show user-contributed content, e.g., Moment Machine [19], and university-related information such as upcoming events and news, public transportation, and the academic calendar [9]. The displays have been in operation every working day from 8:30 to 18:30 since February 2014.

EVALUATION OF THE TOOL

Methodology

We evaluated the tool using a set of depth and webcam videos recorded at the university canteen display installation. In order to evaluate the performance of the tool we used the measures of accuracy, precision, and sensitivity as used by Fawcett [11] and Szeliski [24].

We manually counted and noted the presence and paths of passers-by using the video recording as ground truth. The presence was counted and cross compared by two independent researchers. Then, we checked the output of the tool based on the depth data and compared the results to the ground truth. We counted how many of the people visible in the video were also detected by the tool (True Positives - TP), how many detections by the tool did not correspond to a person (False Positives - FP), and how many manually detected passers-by were not detected by the tool (False Negatives - FN). The measure of accuracy takes into account both false positives and false negatives ($ACC = \frac{TP}{TP+FP+FN}$). Precision, or positive predictive value, indicates the influence of false positives ($PRC = \frac{TP}{TP+FP}$) and sensitivity, or true positive rate, measures the influence of false negatives ($SEN = \frac{TP}{TP+FN}$).

Results

The data-set consisted of 14 recorded videos in a raw (.oni) file format, each of 5 minute duration. Each video contained both depth data and video representation of the scene, as shown in Figure 1 - top right and bottom left. The recorded scenarios were not controlled and we did not interfere with usual behavior of people in front of the display. In total, we manually counted 226 passers-by within the 70 minutes of recordings.

On average in the 14 recorded videos, the accuracy of the tool is 90.53% (min: 78.26%; max: 100.00%; std: 7.46%), sensitivity of the tool is 94.21% (min: 81.81%; max: 100.00%; std: 5.86%) and precision of the tool is 95.97% (min: 85.71%; max: 100.00%; std: 5.75%). An overview of accuracy, sensitivity, and precision for each individual recording is presented in Table 1.

Discussion

Using anonymous three-dimensional representation of the scene, we were able to extract movements and paths of the audience and continuously track them with high levels of accuracy, precision, and sensitivity. However, the main limitations of this approach are due to the characteristics of the depth sensor. Due to the range limitation (up to three meters) and the sensor's view angle (43^o vertical and 57^o horizontal), we were not able to monitor the entire corridor, but only a limited space in front of the display (see Figure 1). Also, due to limited resolution of the sensor (640x480 pixels), the algorithm detected both individuals as multiple persons (e.g., when wearing high-contrast colored clothing), as well as groups of people as individuals (e.g., when moving close together), leading to false positives (FP) and false negatives (FN), as indicated in Table 1. The accuracy of the tool can be increased by deploying multiple cameras [22] or reconstructing occluded parts of the scene [28].

EXAMPLE USE IN EXISTING DISPLAY DEPLOYMENT

The audience monitoring tool helped us better understand the wider situational context of the existing display deployment and its use in front of the university canteen. The tool allowed us to look and visualize the mobility paths at

Table 1 – Accuracy (ACC), Sensitivity (SEN), and Precision (PRC) of the audience monitoring tool

Video	GT	TP	FP	FN	ACC	SEN	PRC
1	40	37	1	3	90.24%	92.50%	97.37%
2	11	11	0	0	100.00%	100.00%	100.00%
3	19	19	3	0	86.36%	100.00%	86.36%
4	11	10	0	1	90.91%	90.91%	100.00%
5	13	11	1	2	78.57%	84.62%	91.67%
6	21	20	2	1	86.96%	95.24%	90.91%
7	24	24	4	0	85.71%	100.00%	85.71%
8	3	3	0	0	100.00%	100.00%	100.00%
9	8	8	0	0	100.00%	100.00%	100.00%
10	17	16	0	1	94.12%	94.12%	100.00%
11	18	16	3	2	76.19%	88.89%	84.21%
12	7	6	0	1	85.71%	85.71%	100.00%
13	9	7	0	2	77.78%	77.78%	100.00%
14	25	23	0	2	92.00%	92.00%	100.00%

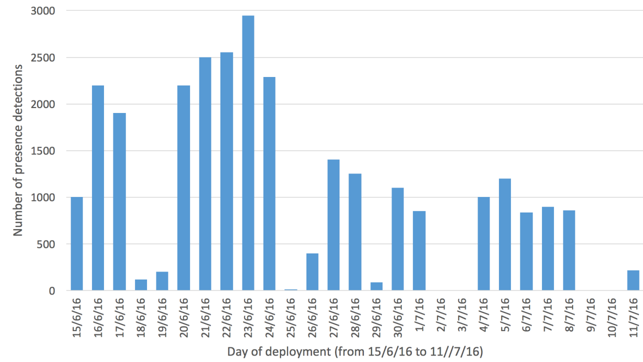

Figure 3 – Example visualization of presence detections per day

specific time instances and obtain insights how people move through the space and approach the display. The mobility information can be coordinated with actual display presentation and touch interactions to reveal display content when passers-by approach the display and start using display applications. This is potentially useful information for evaluating individual display applications and understanding the influence of application and content presentation on the audience behavior.

Also, the tool allowed us to continuously track and visualize audience mobility in the space, even outside the usual display operating hours (8:30 to 18:30 on working days) and better understand how the deployment space is used in general. Estimations of the potential audience size and its usual behavior can help display researchers better plan new display installations, adapt the existing deployments to possible changes in the audience behavior, compare their results to potential use of their displays, and understand any effects of display installations to usual audience mobility.

In order to get general insights about the deployment space, we continuously monitored the space in front of the display located at the entrance of the university canteen for 52 days, from June 15, 2016 to August 05, 2016. In total, we collected 1248 hours of data using only depth information (no video recordings). During this period, the tool detected and tracked 40763 passers-by and display users.

The tool provides visualization of the presence detection over weeks, days, and hours. For example, Figure 3 shows the number of detections over the last two weeks of the

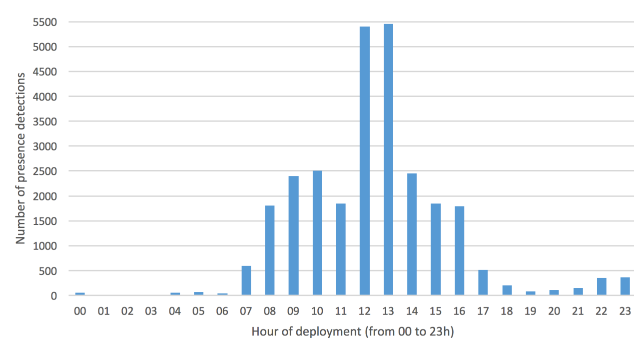

Figure 4 – Example visualization of presence detection per hour

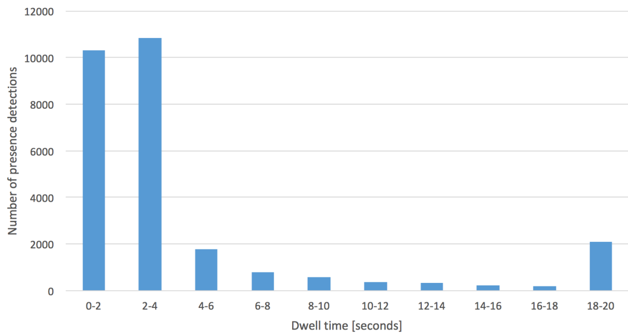

Figure 5 - Example histogram of dwell time in front of display

semester and first two weeks of the summer break. At the beginning of the break, there was still a number of students present due to exam and study sessions. This helped us to convince our university administration to keep the displays in operation during the break when students are around.

Figure 4 shows the sum of presence detections per hour. The lunch time (hours 12 and 13) were the busiest, almost doubling the number of passers-by before and after the lunch. The possibility to quantify the number of passers-by can potentially influence a display business model and content presentation decisions. The presence of people before 8:30 and after 18:30 may indicate a possibility to extend the display presentation hours.

The majority of passers-by stay up to 4 seconds in front of the display as shown in Figure 5. The histogram also indicates that some passers-by spend between 18 and 20 seconds in front of the display. They are either waiting for someone in the space, looking at the display, or interacting with display content. Further coordination with interaction log data can potentially reveal average interaction times with the display and individual display applications.

The tool visualizes the distance of multiple passers-by to the display using a heat map as shown in Figure 6. Capturing heat maps before, during, and after a display deployment can help understand the influence of the deployment to the audience mobility.

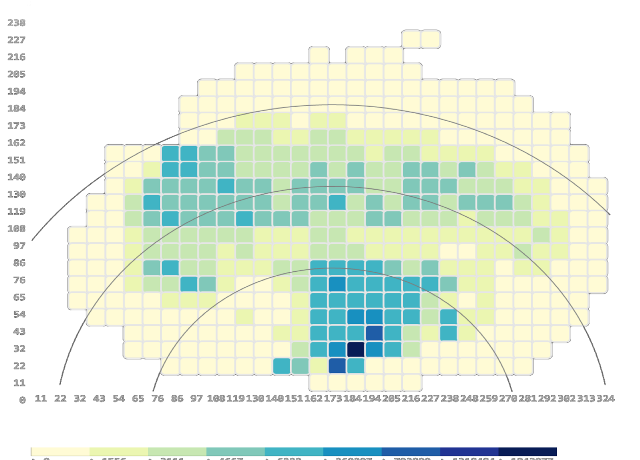

Figure 6 – Heat map representing the distance of passers-by

FUTURE WORK

Analyzing a display deployment using our Audience Monitor helps significantly understanding the display's context and its use. However, this is still a manual process of querying log files for interaction events and using timestamps to adjust the tool's path visualization. We believe that the main potential of the audience behavior information is in providing comprehensive analytics reports by augmenting the existing display data, and enabling opportunities for smart application and content scheduling as we discuss in the following subsections. We are thus currently working on integrating the touch interaction data with the audience monitoring tool and visualizing aggregated information using web analytics and Pheme [17], which we will describe in the next section.

Integration with Web Analytics using Pheme

In order to integrate movement patterns of passers-by with an existing analytics framework, we chose Pheme as a proof of concept to visualize audience tracking data and make it available to other users and stakeholders through existing analytics platforms. While reporting raw analytics events into Pheme is supported by a universal RESTful API, the mapping from our data to third-party analytics engines needs to be explicitly defined.

Web analytics use a limited set of analytics event types to describe interactions and behavior patterns of visitors including *page views* (triggered each time a visitor opens a web site) and generic *events* (allowing a high level of customization). An example analytics report of mapping display data into *page views* and *events* is shown in Figure 7. For the mapping from our datasets to Web analytics, we propose "page views" to represent application and content transitions, and custom "events" to represent audience behavior. In Web analytics, each event can consist of up to four attributes: category, action, label, and value. Typically, the category attributed would be used to support categorization or aggregation of similar events, while actions and labels would describe the individual event that is reported

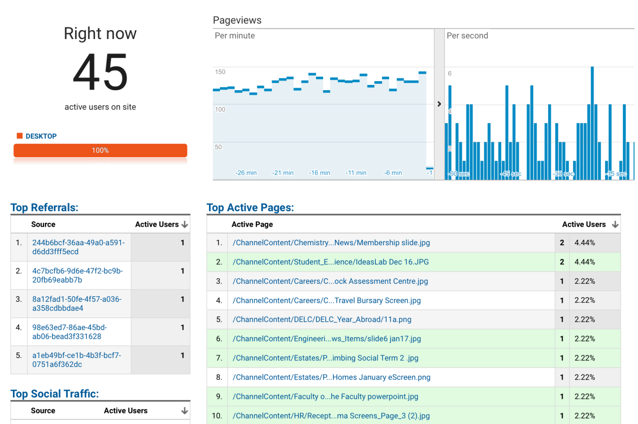

Figure 7 – Example reports generated after mapping and exporting display information to Google Analytics

(e.g. distance to screen). Additional fields can be reported to associate requests with individual users, typically by setting a user and session identifier. This is important for reporting multiple events caused by the same person (e.g. viewer moves in front of the display and different distance measures are reported of the same individual).

Concretely, we propose the following mapping for data reported by our audience behavior tool. As described in [17], the display identifier is mapped onto referrer and client identifier attributes of the analytics system—enabling us to make the source of reports visible in our analytics reports. Due to the higher flexibility and customizability of reporting custom data, we have chosen the *event* type. Table 2 describes a detailed mapping from the datasets we collected to the *events* analytics type. We report the presence, distance, count and dwell time of individuals from the audience as corresponding events. We note that each individual will become a randomized unique identifier assigned so that multiple events caused by the same person are represented and aggregated accordingly. Further, while the count of people and dwell time is aggregated by the Web analytics engine based on the *presence* event (assuming user and session identifiers were set correctly), we have chosen to additionally report these values for verification purposes.

A similar approach can be applied to integrate other types of data such as touch interactions with the displays. The events in the "touch" category could indicate: "coordinates" (x and y coordinates of the touch event), "app change" (when a touch event triggers a change in the app presentation), and "content change" (when users interact with individual content items). Integrating different categories of display relevant data into unified and comprehensive analytics reports enables opportunities for smart application and content scheduling.

Opportunities for Smart Content Scheduling

Integrated and comprehensive reports of display states and audience behavior can help display owners and location managers to better understand their installations and provide opportunities for enabling smarter display behavior and influence application and content scheduling.

Table 2 – Mapping behavioral analytics to Web analytics

Event action	Event value	Description
Presence	Exit / enter	People entering or exiting the scene
Distance	meters	People changing interaction zones or indicate physical distance from the display
Counter	Aggregate number	Number of people in the scene
Dwell time	minutes	Time people spent in front of the display

Digital displays can change their physical power state depending on the presence of passers-by. The displays can extend their operation time and potentially expose content to more users or save power when no users are around. Information and histograms about presence detection by hour, day, and week of deployment can influence business models and content presentation decisions. This information can help quantify and indicate potential sizes of the audience that can view and interact with display content.

Direction, speed, and dwell time of the audience can influence application and content transitions on the display. In our installation, the display has been located between two entrances to the building, classroom area and university canteen. Direction of passers-by can influence the display to present food and drink offers in the canteen (not only during lunch hours), point to upcoming classes and nearby classrooms, or transition to a bus timetable when people move towards the exit of the building. Audience behavior and information about previous touch interactions, or sequences of interactions, can furthermore help the display to identify and present relevant content for different situations.

In our future work, we plan to integrate information generated by analytics reports into novel scheduling approaches for interactive and concurrently running applications [9], thus enabling smarter and more intelligent application and content presentation decisions. Also, we plan to extend the sensing capabilities of the displays using a mobile personalization framework for pervasive displays called Tacita [7]. The framework will help augment audience behavior information with display personalization preferences of both passers-by and display users without compromising their privacy [7].

CONCLUSION

Audience behavior is important aspect when evaluating display installations and understanding a wider situational context of deployments. In this paper we have presented an affordable and easy to install approach for continuous and real-time audience monitoring. In particular, we have shown how the audience monitoring tool can be integrated into an existing display deployment and presented information that such a tool can provide. Using a combination of standard computer vision algorithms and three-dimensional representation of the scene, we were able to achieve high levels of accuracy, precision, and sensitivity. The main potential of the audience behavior data lies in augmenting information about content transitions and presentation times and touch interactions. Using the aggregated data, web based analytics can provide comprehensive and powerful reports. We have made the source code available online for the research community, and believe that this approach will help researchers in evaluating and monitoring existing and future display deployments.

REFERENCES

1. S. Ali and M. Shah, „Floor Fields for Tracking in High Density Crowd Scenes, " In *Proceedings of the 10th European Conference on Computer Vision: Part II*

(ECCV '08), David Forsyth, Philip Torr, and Andrew Zisserman (Eds.). Springer-Verlag, Berlin, Heidelberg, pp 1-14, 2008.

2. Florian Alt, Stefan Schneegaß, Albrecht Schmidt, Jörg Müller, and Nemanja Memarovic. 2012. How to evaluate public displays. In *Proceedings of the 2012 International Symposium on Pervasive Displays* (PerDis '12). ACM, NY, USA. doi:10.1145/2307798.2307815
3. Till Ballendat, Nicolai Marquardt, and Saul Greenberg. 2010. Proxemic interaction: designing for a proximity and orientation-aware environment. In *ACM International Conference on Interactive Tabletops and Surfaces* (ITS '10). ACM, New York, NY, USA, 121-130. doi:10.1145/1936652.1936676
4. Harry Brignull and Yvonne Rogers. Enticing people to interact with large public displays in public spaces. In *In Proceedings of the IFIP International Conference on Human-Computer Interaction (INTERACT 2003)*, pages 17–24, 2003.
5. A. B. Chan, Z. S. J. Liang, N. Vasconcelos, "Privacy preserving crowd monitoring: Counting people without people models or tracking," IEEE Conference on Computer Vision and Pattern Recognition, 2008.
6. N. Davies, M. Langheinrich, R. Jose and A. Schmidt, "Open Display Networks: A Communications Medium for the 21st Century," in *Computer*, vol. 45, no. 5, pp. 58-64, May 2012. doi:10.1109/MC.2012.114
7. Nigel Davies, Marc Langheinrich, Sarah Clinch, Ivan Elhart, Adrian Friday, Thomas Kubitza, and Bholanathsingh Surajbali. 2014. Personalisation and privacy in future pervasive display networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14). ACM, NY, USA, 2357-2366. doi:10.1145/2556288.2557287
8. Ivan Elhart, Marc Langheinrich, Nemanja Memarovic, Elisa Rubegni (2016) A good balance of costs and benefits: convincing a university administration to support the installation of an interactive multi-application display system on campus, Proceedings of the 5th ACM International Symposium on Pervasive Displays - PerDis '16, p. 197-203, New York, New York, USA. doi:10.1145/2914920.2915029
9. Ivan Elhart, Marc Langheinrich, Nemanja Memarovic, Tommi Heikkinen (2014) Scheduling Interactive and Concurrently Running Applications in Pervasive Display Networks, Proceedings of The International Symposium on Pervasive Displays - PerDis '14, p. 104-109, New York, New York, USA. doi:10.1145/2611009.2611039
10. Farinella, G. M., Farioli, G., Battiato, S., Leonardi, S., and Gallo, G. Face re-identification for digital signage applications. In Video Analytics for Audience Measurement (2014), Springer.
11. Tom Fawcett. 2006. An introduction to ROC analysis. *Pattern Recogn. Lett.* 27, 8 (June 2006), 861-874. doi:10.1016/j.patrec.2005.10.010
12. Fraunhofer IIS. Fraunhofer AVARD. http://www.iis.fraunhofer.de/en/ff/bsy/tech/bildanalyse/avard.html
13. Nicholas Gillian, Sara Pfenninger, Spencer Russell, and Joseph A. Paradiso. 2014. Gestures Everywhere: A Multimodal Sensor Fusion and Analysis Framework for Pervasive Displays. In *Proceedings of The International Symposium on Pervasive Displays* (PerDis '14), Sven Gehring (Ed.). ACM, New York, NY, USA. doi:10.1145/2611009.2611032
14. Intel Corporation. Intel AIM Suite. https://aimsuite.intel.com/
15. Michelis, D., and Müller, J. 2011. The Audience Funnel: Observations of Gesture Based Interaction With Multiple Large Displays in a City Center. *Int. J. of HCI 27*, 6 (May 2011), 562 - 579. doi:10.1080/10447318.2011.555299
16. Shahram Izadi, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Dustin Freeman, Andrew Davison, and Andrew Fitzgibbon. 2011. KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera. In *Proceedings of the 24th annual ACM symposium on User interface software and technology* (UIST '11). ACM, NY, USA, 559-568. doi:10.1145/2047196.2047270
17. M. Mikusz, S. Clinch, R. Jones, M. Harding, C. Winstanley and N. Davies, "Repurposing Web Analytics to Support the IoT," in *Computer*, vol. 48, no. 9, pp. 42-49, Sept. 2015. doi:10.1109/MC.2015.260
18. Nemanja Memarovic, Ivan Elhart, Marc Langheinrich (2011) FunSquare: first experiences with autopoiesic content, Proceedings of the 10th International Conference on Mobile and Ubiquitous Multimedia - MUM '11, p. 175-184, New York, New York, USA: ACM Press, doi:10.1145/2107596.2107619
19. Nemanja Memarovic, Ivan Elhart, and Elisa Rubegni. 2016. "Fun place within a serious space": stimulating community interaction and engagement through situated snapshots in a university setting. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia* (MUM '16). ACM. doi:10.1145/3012709.3012710
20. Jörg Müller, Juliane Exeler, Markus Buzeck, and Antonio Krüger. 2009. ReflectiveSigns: Digital Signs That Adapt to Audience Attention. In *Proceedings of the 7th International Conference on Pervasive Computing* (Pervasive '09), Hideyuki Tokuda, Michael Beigl, Adrian Friday, A. J. Brush, and Yoshito Tobe (Eds.). Springer-Verlag, Berlin, Heidelberg, 17-24. doi:10.1007/978-3-642-01516-8_3

21. Timo Ojala, Hannu Kukka, Tomas Lindén, Tommi Heikkinen, Marko Jurmu, Simo Hosio, and Fabio Kruger. 2010. UBI-Hotspot 1.0: Large-Scale Long-Term Deployment of Interactive Public Displays in a City Center. In *Proceedings of the 2010 Fifth International Conference on Internet and Web Applications and Services* (ICIW '10), Washington, DC, USA, 285-294. doi:10.1109/ICIW.2010.49
22. Stefan Seer, Norbert Brändle, Carlo Ratti. 2014. Kinects and human kinetics: A new approach for studying pedestrian behavior. In Transportation Research Part C: Emerging Technologies, Volume 48, November 2014, 212-228. doi:10.1016/j.trc.2014.08.012.
23. Jiamin Shi and Florian Alt. 2016. The Anonymous Audience Analyzer: Visualizing Audience Behavior in Public Space. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (CHI EA '16). ACM, New York, NY, USA, 3766-3769. doi:10.1145/2851581.2890256
24. Richard Szeliski. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010. doi:10.1007/978-1-84882-935-0
25. Phil Tian, Addicam V. Sanjay, Kunapareddy Chiranjeevi, and Shahzad Malik Malik. 2012. Intelligent advertising framework for digital signage. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining* (KDD '12). ACM, New York, NY, USA, 1532-1535. doi:10.1145/2339530.2339773
26. Daniel Vogel and Ravin Balakrishnan. 2004. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the 17th annual ACM symposium on User interface software and technology* (UIST '04). ACM, New York, NY, USA, 137-146. doi:10.1145/1029632.1029656
27. Julie R. Williamson and John Williamson. 2014. Analysing Pedestrian Traffic Around Public Displays. In *Proceedings of the International Symposium on Pervasive Displays* (PerDis '14), ACM, NY, USA. doi:10.1145/2611009.2611022
28. Chi-Jui Wu, Aaron Quigley and David Harris-Birtill. 2017. Personal and Ubiquitous Computing (2017) 21: 125. doi:10.1007/s00779-016-0997-6