Pre-trip Training System for Seniors and People with Disabilities using Annotated Panoramic Video

Hao Dong and Aura Ganz
Electrical and Computer Engineering, University of Massachusetts, 151 Holdsworth Way, 01003, Amherst MA, U.S.A.

Keywords: Pre-trip Training, Panoramic Video, Transportation.

Abstract: This paper presents a scalable and user-friendly pre-trip training system for seniors and people with disabilities using panoramic videos. The proposed system allows travel trainers to annotate the videos according to the user disability and requirements. Such annotations will be displayed to the users during the training process. After training with the system, seniors and people with disabilities will be more likely to choose fixed route services while traveling in complex subway systems and indoor transportation hubs. Therefore, the use of the proposed platform will result in significant savings of paratransit services.

1 INTRODUCTION

According to the US Census Bureau (Census.gov 2016) 56.7 million people (19% of the US population) – had a disability in 2010. Moreover, by 2040 Americans aged 65 or older increase from 14.5% to 21.7% of the population (Aoa.acl.gov 2016). Subway stations and transportation hubs include complex multi-story underground buildings with crowded and noisy environments, which can overwhelm seniors and people with disabilities. Urban residents including the disabled and mobility-impaired elderly are offered paratransit services as required by 1990 Americans with Disabilities Act mandate.

According to a recent report (Kaufman et al., 2016), paratransit demand is growing nationwide and costs continually increase (now $5.2 billion nationwide). In New York City, paratransit serves 144,000 subscribers at $456 million per year; in the Chicago region, 50,000 subscribers are served at $137 million per year; in Boston, 80,000 at $75 million per year.

One effort to reduce the paratransit cost is to provide more accessibility in fixed route public transportation systems. Considering the users' disabilities and the complexity of these indoor transportation environments, travel trainers are assigned to prepare them to travel independently. However, since the training budgets are limited, many seniors and people with disabilities will not be exposed to such valuable training.

This paper attempts to reduce the cost of paratransit services and enhance the confidence of seniors and people with disabilities to use fixed route transportation. We introduce a virtual pre-trip training system that enables them to get familiar with the structure and features of the subway station or transportation hub. Such familiarity will instil confidence in these users and make their travel experience safer and more efficient. Therefore, they will be inclined to use fixed route services more frequently instead of the high cost paratransit services. Moreover, we provide tools for travel trainers that can integrate their expertize in the proposed system enabling users to train at their own pace, at their chosen time and from their own home.

Different from traditional training, which requires both of trainers and trainees to be present in the target building (e.g. subway station), the virtual training system uses panoramic videos to represent the target environment. In such a virtual environment, trainers can annotate the video with the necessary information tailored to the user's disability or requirements. The annotations will be shown to the users when they are relevant to the training context.

The system includes the following parts:

- **Video Recording:** In order to generate panoramic videos, 4 GoPro cameras are used to capture the environment in four main directions, i.e. front, back, left and right. The

Dong, H. and Ganz, A.
Pre-trip Training System for Seniors and People with Disabilities using Annotated Panoramic Video.
DOI: 10.5220/0006312201500156
In *Proceedings of the 3rd International Conference on Information and Communication Technologies for Ageing Well and e-Health (ICT4AWE 2017)*, pages 150-156
ISBN: 978-989-758-251-6
Copyright © 2017 by SCITEPRESS – Science and Technology Publications, Lda. All rights reserved

personnel will capture the video in the pre-planned paths in the target building. Using these videos we generate a geo-referenced panoramic video.

- **Annotations:** The trainer will annotate the videos to meet different requirements for people with different disabilities. These annotations can include any travel information that can benefit the user, such as landmarks, facility information, or even notes.
- **Virtual Training:** Seniors and people with disabilities can use this system in three different modes: 1) take a virtual tour of the paths selected by their trainers, 2) visit a specific landmark, or 3) explore the building by themselves.

The paper is organized as follows. In the next section we introduce related work and in Section 3 we present the system architecture. Section 4 presents a case study of how this system will be used and Section 5 concludes the paper.

2 RELATED WORK

Constructivism (Duffy and Jonassen 2013) is a philosophical viewpoint that students can construct their knowledge and understanding in a contextual and visually rich environment by interacting with the training information. Many researchers and game developers start to design games or simulators for different training purpose other than pure entertainment. This type of games is called “serious games”, which are described as the next wave of technology-mediated learning. A well-known example is Microsoft Flight Simulator, a comprehensive simulation of civil aviation (Microsoft.com 2016). There are also a number of projects focusing on special training for people with disabilities, such as training for mobility and navigation skills for visually impaired children (Allain 2015) (Simões 2014) (Magnusson 2011) (Cavaco 2015), and cognitive training and screening for Alzheimer patients (Bouchard 2012) (Boletsis 2016) (Imbeault 2011) (Manera 2015).

It is well known that the quality of the game environment can determine user’s satisfaction. For instance, the system presented in (Sánchez 2010), Audio-based Environment Simulator (AbES), only constructs a 2D tile-based environment, as the game aims to provide projected sound to visually impaired and blind users based on 2D spatial relationships. XVR, which is an Emergency Training Platform, builds very vivid 3D models and environments to recover from stress in disaster field (Xvrsim.com 2016). The construction of the 2D environment built in AbES is easy to construct since it includes less details of the target building. However, modelling the environment used in XVR will be very time-consuming.

In this paper we propose a training system that uses geo-referenced panoramic videos that represent the virtual environment. Using panoramic videos has several advantages compared with 2D or 3D modelling environments presented above. First, the virtual environment represented by the panoramic video can provide a similar experience as walking through the real environment, including obstacles, furniture, decorations, and even noise and crowd.

Second, the virtual environment preparation obtained by recording videos in the target building is a simpler process than generating a 3D environment. The video recording in the target building requires lower level skills than modelling a 3D building structure from a blueprint.

There are a few systems that use videos to generate tourist information guides. The system described in (Mildner 2013) uses multiple video sequences to generate a virtual video tour of an outdoor environment. In (Zhang 2010) the authors present a novel system for registering videos. Instead of using video sequences, given start and end points, the system in (Peng 2010) can automatically connect to Google Maps to query Street View pictures in the planned route, and generate a smooth scenic video. In (Zhao 2015) the authors propose to use video captured by a dashboard camera to construct a city virtual tour. The viewing and interaction in an emerging type of interactive TV explored in (Zoric 2013) are showing the benefits of interacting with panoramic content. Streaming a panoramic video on mobile devices is also attempted in (Barkhuus 2014). To the best of the authors’ knowledge there are no published systems that consider pre-trip planning systems in indoor environments for seniors and people with disabilities.

3 SYSTEM ARCHITECTURE

The system architecture, which is shown in Figure 1, includes four components: the video recording process, the server, the annotation application, and the training application.

Both training and annotation applications are developed using Unity3D, which is 3D game development engine. The annotation application is a desktop application designed for trainers to add and edit annotations in the panoramic video. The trainers will view the panoramic video in a 3D renderer display, and then edit any necessary trainee information through the user interface. All annotations will be uploaded to the server and saved in the annotation database. The training application enables the trainees to view the panoramic video in a 3D renderer display with overlaid annotation. Users can view a selected video directly, or video sequences by selecting the source and destination of a path.

We provide a brief description of each component below.

3.1 Video Recording Process

To record the target environment, the video recording process uses a helmet-mounted rig (Figure 2) with 4 GoPro cameras. All videos will be uploaded to the server. Similar to GIS representation of maps (Data.geocomm.com, 2016), we generate a graph of our indoor environment using its Blueprint. All the links in this graph will be recorded in both directions.

The GoPro camera rig we use for video recording evenly positions 4 cameras to cover the recording of all four directions on the horizon plane. The panoramic video generated using this layout is displayed in Figure 3. The black areas on the top and bottom parts indicate blank top view and bottom view, which are not informative in indoor environments.

In order to show this video correctly, we apply it as texture on a spherical shape, which is centred at the user's camera. Simply panning the camera along with user's operations can simulate turn movements. By playing the video, the user can watch any direction along the path.

3.2 Server

The server functionality includes:

- **Video Reception and Storage:** the server stores the videos captured by the video recording process in a video file system.
- **Generate a Panoramic Video with 360-Degree Panning View:** To obtain the panoramic video we synchronize and stitched up together the videos based on common features detected in overlapped areas captured by adjacent cameras. The panoramic videos are stored in another video file system.
- **Content Loading Services**: One service selects and transmits content for annotation; the other service displays pre-trip training annotated video.

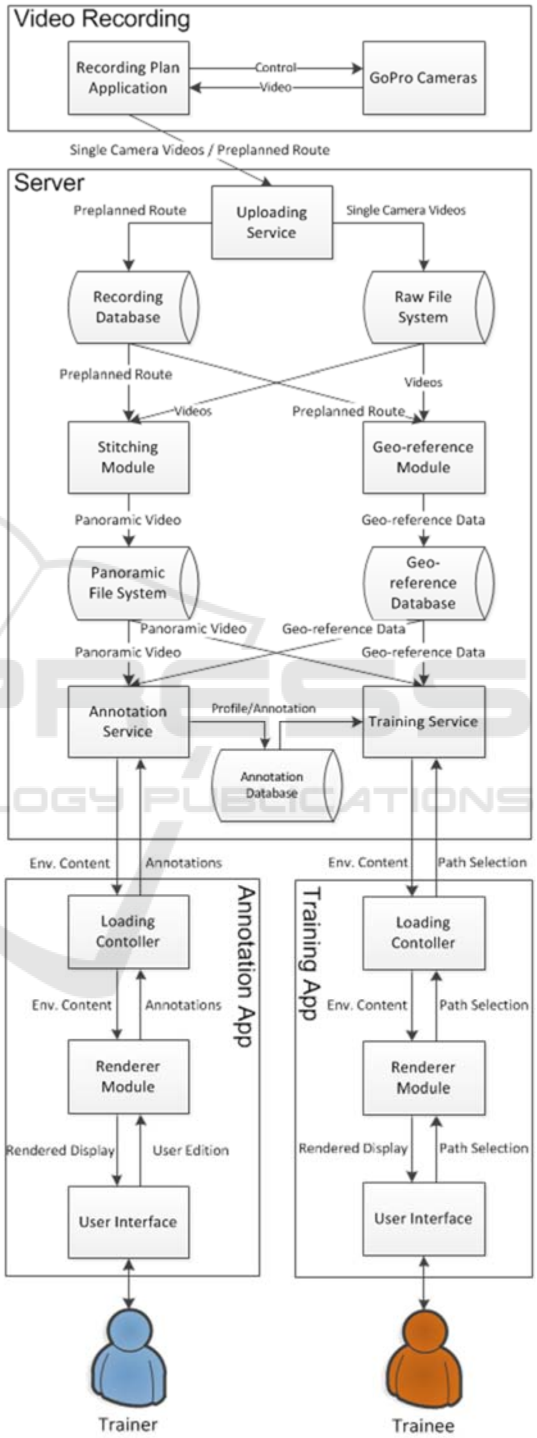

Figure 1: System Architecture.

3.3 Annotation Application

Two types of annotations can be added into the videos. One indicates landmarks in distance, which requires the trainer to edit the red beam that points to them. The other annotation type includes landmarks, which indicate a location or area in the current position. For example, in a subway station the trainers can annotate a fare gate. For visually impaired, the audio annotation can be “The fare gate with a beeping sound is located in front of you”. For cognitive impaired, the text and/or audio annotation can be “ The fare gate is located under the green light in front of you”.

These annotations will also be used for destination selection and wayfinding algorithm. When the annotation process is completed, the trainer can click on the “export content” button to synchronize with the server.

The trainer will determine paths (each path is defined by two waypoints) and/or specific landmarks that the user needs to explore using the training application. For example, the trainer will generate tasks pertinent to emergency evacuation. Such tasks may include multiple sources leading to an accessible exit as well as designate specific exits as landmarks to further explore.

3.4 Training Application

The training application is designed to represent the virtual environment including the trainer’s annotations. Users can start a training session from a specific landmark of interest or the start point of a selected path, and control their movements using the keyboard to “proceed”, “look left”, and “look right”. Annotations will be shown when the user’s position is within a certain distance from the landmark.

4 CASE STUDY

In this case study we introduce the system deployment and usage in a subway section of North Station, Boston, MA. We introduce the following steps: video recording, annotation application, and training application.

4.1 Video Recording

Using the Blueprints, we first plan the recording paths that cover the most utilized paths between different locations of interest, such as entrances, ticket machines, ticket gates, and platforms. We

Figure 2: Camera Helmet with 4 GoPro Cameras.

Figure 3: Example Frame of Panoramic Video.

record the videos by following these paths wearing the camera rig described in the previous section and manually record key waypoints in each path like start locations, end locations, and turning locations. After the recording is finished, the video sequences and waypoints of the associated paths will be uploaded to the server for stitching and geo-referencing.

4.2 Annotation Design

We assume that the trainer will prepare the application for seniors. The trainer will first select the building and profile using the interface shown in Figure 4a. Then the annotation application will load North Station Subway Station from the server and show all available video clips in a list (shown in Figure 4b). The trainer will choose and play a video from the list (shown in Figure 4c) and select important landmarks, e.g. the escalator connecting to the platform of another subway line. Through the annotation interface shown in Figure 4d the trainer adds relevant information to each landmark. The red beam shown at bottom center indicates the direction of this landmark relative to the position of the current frame.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

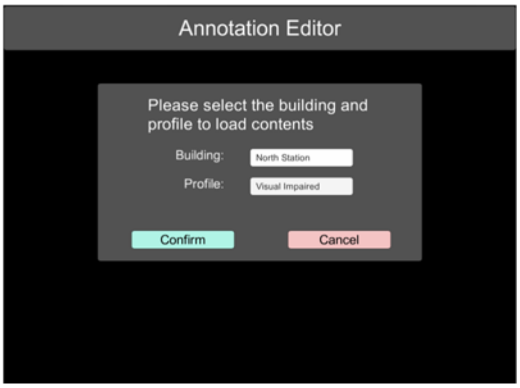

Figure 4a: Screenshot of building and profile selection.

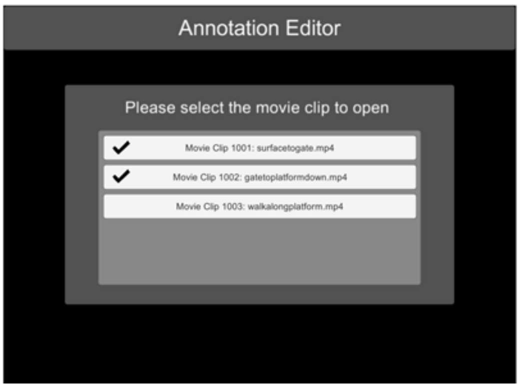

Figure 4b: Screenshot of Video Material Selection.

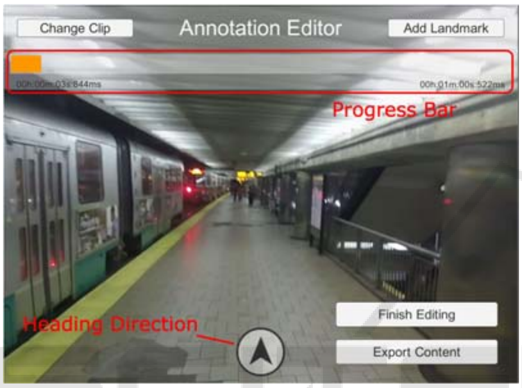

Figure 4c: Screenshot of panoramic video rendering.

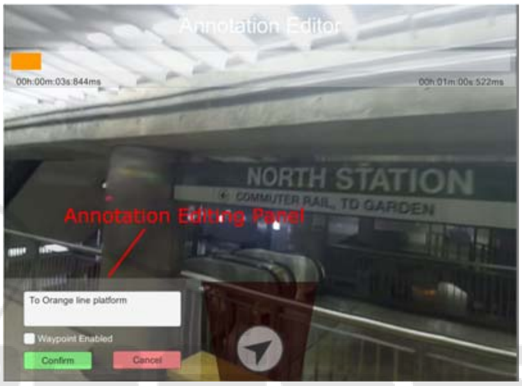

Figure 4d: Screenshot of annotation editing.

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS

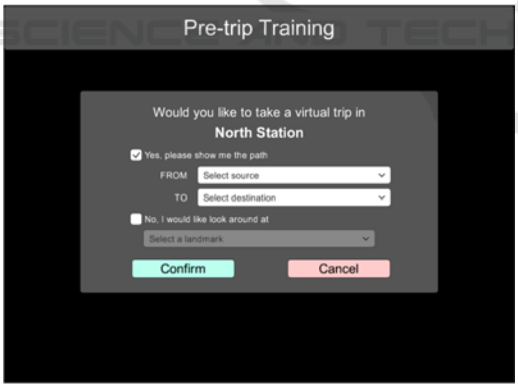

Figure 5a: Screenshot of path and landmark selection.

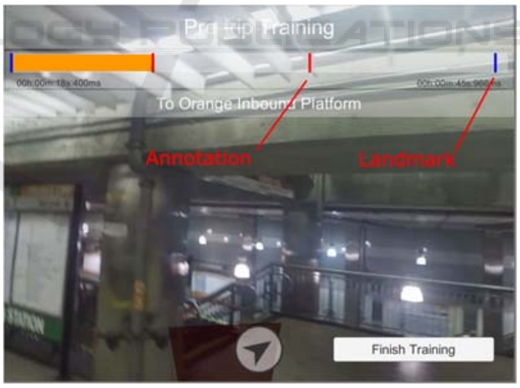

Figure 5b: Screenshot of training view.

4.3 Training Application

Following the tasks assigned by the trainer, the trainee explores North Station using the training application (Figure 5). After selecting the building name and his/her profile (Figure 4a), the trainee can either select a path (mention source and destination) or select a landmark (Figure 5a) to start the training session. When the trainee gets close to a landmark, a red “beam” will be overlaid and pointing to it (Figure 5b) and a description of this landmark will be shown on top.

5 CONCLUSIONS AND FUTURE WORK

The authors introduced a pre-trip training platform for seniors and people with disabilities, which use panoramic video to represent the physical environment. The travel trainers can create annotations of important travel information for training purpose. This platform can increase the likelihood that seniors and people with disabilities will use fixed route services instead of paratransit services, significantly reducing the paratransit cost. Our next steps are to start trials with seniors and people with disabilities and understand in depth how to optimize this platform in order to provide maximum benefits to these users.

REFERENCES

Allain, K., Dado, B., Van Gelderen, M., Hokke, O., Oliveira, M., Bidarra, R., Gaubitch, N.D., Hendriks, R.C. and Kybartas, B., 2015, March. An audio game for training navigation skills of blind children. In *Sonic Interactions for Virtual Environments (SIVE), 2015 IEEE 2nd VR* Workshop on (pp. 1-4). IEEE.

Aoa.acl.gov. 2016. Aging Statistic. [online] Available at: https://aoa.acl.gov/Aging_Statistics/Index.aspx. [Accessed 23 Feb. 2017]

Barkhuus, L., Engstrom, A., and Zoric, G., 2014. Watching the footwork: Second screen interaction at a dance and music performance. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems* (pp. 1305-1314), ACM.

Boletsis, C., & McCallum, S. 2016. Smartkuber: A Serious Game for Cognitive Health Screening of Elderly Players. Games for health journal.

Bouchard, B., Imbeault, F., Bouzouane, A., & Menelas, B. A. J. 2012, September. Developing serious games specifically adapted to people suffering from Alzheimer. In *International Conference on Serious Games Development and Applications* (pp. 243-254). Springer Berlin Heidelberg.

Cavaco, S., Simões, D., & Silva, T. 2015, November. Spatialized audio in a vision rehabilitation game for training orientation and mobility skills. In *Proceedings of the 18th International Conference on Digital Audio Effects (DAFx-15),* NTNU.

Census.gov. 2016. Nearly 1 in 5 People Have a Disability in the U.S. [online] Available at: https://www.census.gov/newsroom/releases/archives/miscellaneous/cb12-134.html. [Accessed 25 Nov. 2016].

Data.geocomm.com. (2016). Free GIS Data - GIS Data Depot. [online] Available at: http://data.geocomm.com/helpdesk/formats.html [Accessed 27 Nov. 2016].

Duffy, T.M. and Jonassen, D.H. eds., 2013. *Constructivism and the technology of instruction: A conversation.* Routledge.

Goodwill, J.A. and Carapella, H., 2008. Creative ways to manage paratransit costs. Center for Urban Transportation Research University of South Florida.

Imbeault, F., Bouchard, B., & Bouzouane, A. 2011, November. Serious games in cognitive training for Alzheimer's patients. In *Serious Games and Applications for Health (SeGAH), 2011* IEEE 1st International Conference on(pp. 1-8). IEEE.

Kaufman, S.M., Smith, A., O'Connell, J., Marulli, D., 2016. Intelligent Paratransit. NYU Rudin Center for Transportation. [online] Available at: http://wagner.nyu.edu/rudincenter/wp-content/uploads/2016/09/INTELLIGENT_PARATRANSIT.pdf. [Accessed 23 Feb. 2017].

Magnusson, C., Waern, A., Gröhn, K.R., Bjernryd, Å., Bernhardsson, H., Jakobsson, A., Salo, J., Wallon, M. and Hedvall, P.O., 2011, August. Navigating the world and learning to like it: mobility training through a pervasive game. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services* (pp. 285-294). ACM.

Manera, V., Petit, P. D., Derreumaux, A., Orvieto, I., Romagnoli, M., Lyttle, G., ... & Robert, P. H. 2015. 'Kitchen and cooking,'a serious game for mild cognitive impairment and Alzheimer's disease: a pilot study. Frontiers in aging neuroscience, 7, 24.

Microsoft.com. (2016). Product Information. [online]. Available: https://www.microsoft.com/Products/Games/FSInsider/product/Pages/. [Accessed: 30- Nov- 2016].

Mildner, P., Claus, F., Kopf, S., & Effelsberg, W. 2013, February. Navigating videos by location. In *Proceedings of the 5th Workshop on Mobile Video* (pp. 43-48). ACM.

Peng, C., Chen, B. Y., & Tsai, C. H. 2010, December. Integrated google maps and smooth street view videos for route planning. In *Computer Symposium (ICS), 2010* International (pp. 319-324). IEEE.

Rita.dot.gov. 2016. Data Analysis. [online] Available at: http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/freedom_to_travel/html/data_analysis.html [Accessed 25 Nov. 2016].

Sánchez, J., Sáenz, M., Pascual-Leone, A. and Merabet, L., 2010, April. Enhancing navigation skills through audio gaming. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems* (pp. 3991-3996). ACM.

Simões, D. and Cavaco, S., 2014, November. An orientation game with 3D spatialized audio for visually impaired children. In *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology* (p. 37). ACM.

Web.mta.info. 2016. Subways. [online] Available at: http://web.mta.info/nyct/facts/ffsubway.htm [Accessed 25 Nov. 2016].

Xvrsim.com. (2016). Virtual Reality training software for safety and security. [online]. Available: http://www.xvrsim.com/. [Accessed: 30- Nov- 2016].

Zhang, B., Li, Q., Chao, H., Chen, B., Ofek, E., & Xu, Y. Q. 2010, November. Annotating and navigating tourist videos. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic*

Information Systems (pp. 260-269). ACM.

Zhao, G., Zhang, M., Li, T., Chen, S. C., & Rishe, N. 2015, August. City recorder: Virtual city tour using geo-referenced videos. In *Information Reuse and Integration (IRI), 2015* IEEE International Conference on (pp. 281-286). IEEE.

Zoric, G., Barkhuus, L., Engstrom, A., and Onnevall, E., 2013. Panoramic video: Design challenges and implications for content interaction. In *Proceedings of the 11th European Conference on Interactive TV and Video* (pp. 153-162), ACM.

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS