

SONA: Improving Situational Awareness of Geotagged Information using Tangible Interfaces

Seung Youb Ssin, Joanne E. Zucco, James A. Walsh, Ross T. Smith, Bruce H. Thomas

Wearable Computer Lab
University of South Australia

Mawson Lakes, South Australia, Australia

{Seung_Youb.Ssin, Joanne.Zucco, James.Walsh, Ross.Smith, Bruce.Thomas}@unisa.edu.au



Fig. 1. The SONA system components: (a) SONA environment overview, (b) Users operating the SONA system with TUI, (c) Tangible camera control input device, and (d) A virtual city presented on the SONA system.

Abstract—This paper introduces SONA (Spatiotemporal system Organized for Natural Analysis), a tabletop and tangible controller system for exploring geotagged information, and more specifically, CCTV. SONA’s goal is to support a more natural method of interacting with data. Our new interactions are placed in the context of a physical security environment, closed circuit television (CCTV). We present a three-layered detail on demand set of view filters for CCTV feeds on a digital map. These filters are controlled with a novel tangible device for direct interaction. We validate SONA’s tangible controller approach with a user study comparing SONA with the existing CCTV multi-screen method. The results of the study show that SONA’s tangible interaction method is superior to the multi-screen approach, both in terms of quantitative results, and is preferred by users.

Keywords— *Tangible User Interface; Geo-visualization; Focus plus context visualization; Multiple level/layered view; Sensor tracking; Natural User Interface; Surveillance/Security Cameras; CCTV; Tabletop; Optitrack; 3D virtual city*

I. INTRODUCTION

Urban-focused geospatial data is collected and utilized across a wide range of fields [1, 2], including policing and crisis management [3, 4]. A compelling real-world example is the pervasive nature of real-time closed circuit television (CCTV) systems to monitor urban environments. Monitoring traffic conditions, locating a person of interest, security assessments, and disaster management are examples of situations where operators are required to monitor, respond, and report information quickly. CCTV systems commonly employ mouse, keyboard, and joystick inputs to navigate video feeds. Such systems are often equipped with manual zoom, pan, and tilt operations.

Fig. 2 depicts a conventional CCTV monitoring system employing a joystick, keyboard, and mouse as input devices. Enhancing such a system with the notion of "perceptible affordances" [5], we can create links between perception and goal-oriented actions for a more direct means of exploring and controlling the system.



Fig. 2. A traditional CCTV center [6].

This work leverages perceptible affordances and direct interaction by incorporating Tangible User Interfaces (TUIs) with a tabletop display [7], aiming to enable a natural method of exploring geotagged information feeds. We present the *Spatiotemporal system Organized for Natural Analysis* (SONA) (Fig. 1), a tangibly controlled tabletop system for exploring geotagged media feeds, and more specifically CCTV feeds (recorded or streaming). SONA aims to increase a user’s situational awareness and understanding by presenting data in its geotagged location, and by providing natural and direct manipulation controls. SONA provides a three-layered view based on the source geographic data. The first layer of the visualization is the event view. Extending the concepts of Anwar et al. [8], the event view is used to express locations of interest on a map (e.g. accidents/incidents). The events are generic and could represent traffic accidents, congestion, or any other geotagged event, and as such, the detection of these generic events is outside the focus of this work.

Upon selecting an event, the second layer presents the location of geotagged data sources. In the context of the CCTV domain, these are camera locations. The third layer is a data layer used to view the actual content of the data source. We extended Anwar et al.’s event notification system with a focus plus context approach, in addition to incorporating live video data streams [8].

Limitations of current CCTV monitoring systems and other geotagged data include the following: a) the exploration of large-scale data streams is difficult to control and monitor without specialized training and b) the difficulty in matching data to its geographic source location given multiple feeds. SONA overcomes these limitations by employing a tabletop system [9], a combination of novel interfaces (using in-situ/geotagged camera feeds and TUIs), along with a three-layered view to present the data in geographical context.

SONA was evaluated against the existing approach used in large scale CCTV monitoring setups. The contributions of this research are as follows:

- a novel five degrees-of-freedom tabletop TUI for exploring CCTV environments and
- a user study validating the use of the novel TUI.

The following section introduces related work, followed by a detailed description of SONA. The remaining sections present an evaluation of SONA's functionality by means of a user study and a discussion of the outcomes. The last section provides concluding remarks and future research directions.

II. RELATED WORK

SONA covers several different areas of research, including aspects of geographic visualization and tangible interfaces. While not the core focus of SONA, our new system utilizes geospatial visualization techniques to explore camera feeds as they relate to events occurring on the map. The drill down approach used within SONA is a direct application of the *Visual Information Seeking Mantra* formalized by Shneiderman [10], adapted by countless other systems, such as *FingerGlass* [11], which utilized a three-layered visualization system for interacting with generic data.

One interesting application of such an approach in exploring geotagged events was *Nu-View* [12], a co-located collaboration system for medical experts to track the spread of diseases. In this vein and similar to SONA, Guerlain, Cortina, and Renault [13] presented *Smart City Logistics*, a tabletop Geographical Information System (GIS) system with tangible controllers for collective decision making, focusing specifically on freight planning in European cities. Also in this area, Anwar, Nagel, and Ratti [8] introduced *Traffic Origins* to represent road incidents over time. Events are visualized by an expanding circle of color which can then be interacted with. This approach is further extended and utilized in SONA as the first view of the three-layered view.

With a similar focus on GIS, Ying [14] constructed the *Three-dimensional Geographic Information System* (3D GIS), a 3D city model for the design and maintenance of CCTV systems using virtual cities. We utilize a similar approach of virtual cameras in a preconfigured virtual city for our evaluation. In exploring cities tangibly, Alaçam et al. [15] expressed 3D perspective views on a 2D map in the *Tangible Desktop* project.

In the tangibles space, Fitzmaurice, Ishii, and Buxton defined *Graspable User Interfaces* [16], with Fitzmaurice and Buxton [17] finding benefits in the performance of space-multiplexed inputs (tangibles), compared to time-multiplexed inputs (traditional UIs). Walsh, et al. found tangible user interfaces were effective for technical and non-technical users [18]. In looking at tangibles for spatial tasks, Qin et al. [3] implemented the *uEmergency* tabletop system, to support an emergency center environment where people can work in a situation of urgency. They demonstrated how a tabletop system could be used to build an interactive environment where the geographical view can be used to improve on-site first aid. The overlap of such systems as *Nu-View* or *Traffic Origins* has not been explored and is a primary area of focus for SONA.

Mateu, Lasala, and Alamán [19] used mixed reality with a Kinect to develop *OpenSim*, combining virtual objects with tangible interfaces to teach high school students developing physical activity and social skills. The intuitive nature of using simple, tangible objects to enhanced participants' task comprehension when using the mixed reality education system. However, more research is required for appropriate techniques for the sophisticated manipulation of a TUI for SONA operations designed to operate with a 3 level visualization.

Exploring the geospatial space and using tangible interactions, Jones and Maquil [20] composed a learning interface using tangible objects with a tabletop based on a geospatial system. The interactions with the TUIs were zoom, pan, adding layers, and working with layers of geospatial graphics. In a similar vein, the work presented in this paper introduces a method of controlling CCTV cameras without special training and synchronized camera shaped TUIs.

III. SONA

SONA is a novel tool for the presentation and exploration of geotagged streaming information in context. This ability allows novice users and experts to control and explore a dataset without any specific training. Whilst the underlying approaches presented in this paper can be applied to other any geotagged data, we focus on exploring CCTV to provide motivation and context. CCTV installations generate large quantities of information from multiple locations. The current multiple screen (Fig. 2) approach not only removes the location context from the camera's feed but also has the potential to overwhelm the user with information. In particular, as the number of CCTV feeds is increased to monitor a facility, the task of monitoring and controlling the camera direction is also more difficult.

A. Hardware

SONA's hardware (Fig 1) consists of an 116x84x97cm table with an inlaid 42inch display. Natural Point *OptiTrack*¹ cameras are mounted around tabletop to 6DOF track the TUIs, and a Kinect is mounted on the ceiling facing the table to enable touch interaction. A purpose-built touch screen could also be used as an alternative to the Kinect.

¹ <http://www.optitrack.com>

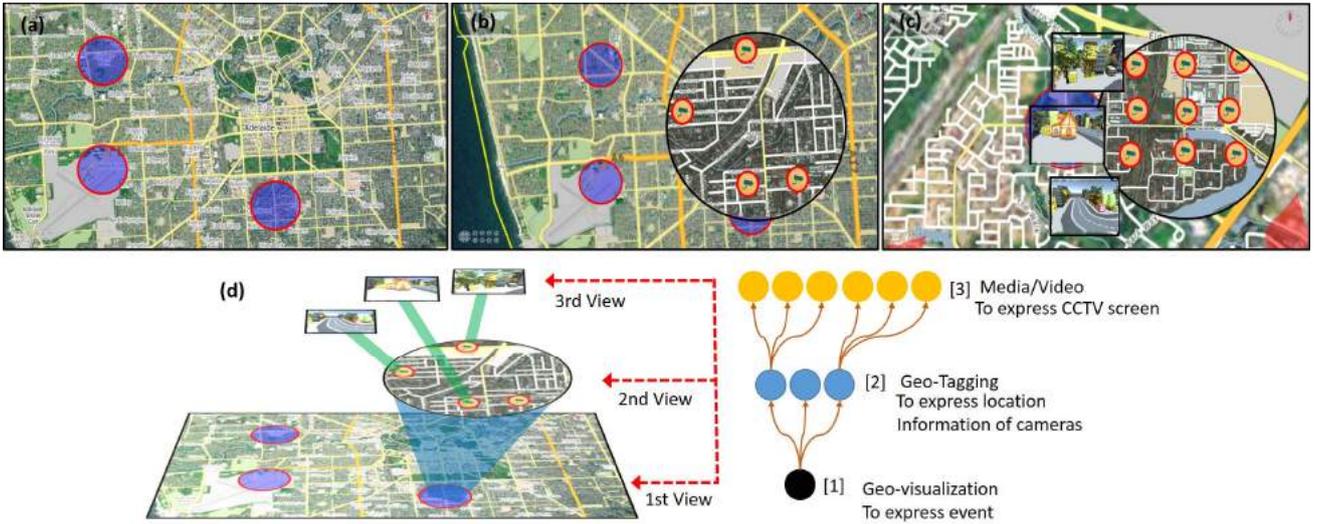


Fig. 3. Three-layered view: (a) the first layer event view, (b) the overlapped second layer information view, (c) the third layered CCTV view, and (d) the structure of three-layered visualization.

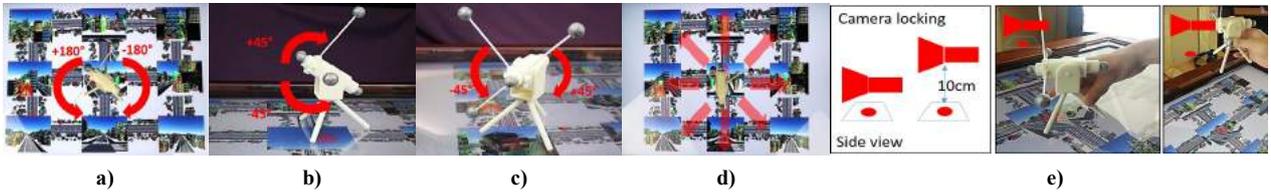


Fig. 4. a) Panning with the TUI. b) Tilting with the TUI. c) Zooming with the TUI. d) TUI can be moved across the display to select a camera. e) Locking the TUI to a camera.

B. 3 Layer Visualization Interface

As highlighted previously, SONA provides a three-layer approach to navigating the camera feeds in context with a map. The first layered view of SONA is the events view (Fig. 3a), and this view represents the position of the events (such as incidents, accidents, disasters, traffic jam, etc.) occurring as blue circles on the map. If users touch the circle, the second layer view opens. The second layer view shows the locations of all CCTV cameras in the surrounding area (Fig. 3b). If users touch one of the camera's icons, the CCTV feed displays (Fig. 3c). Given the example use case of SONA (CCTV feeds), data sources are represented by camera icons on a map reflecting their real locations. When a user touches the geotagged data, the CCTV feed corresponding to the position is opened alongside the icon. Touching them again closes them.

Existing CCTV systems serve to show information from as many sources as possible [6]. However, as shown in Fig. 2, information is presented without geo-referencing video feed data over multiple screens. We theorize that the situational awareness and understanding of spatial layout from the existing systems is difficult for the user and might be improved by leveraging a tabletop display, TUI and 3-layer visualization technique.

As such, SONA displays CCTV videos in their respective locations, allowing feeds to still be shown simultaneously while maintaining contextual information with the placement (Fig. 3d). Using this approach, users can identify an event of interest, identify the location of interest, and drill down to the actual event from the camera streams. The aim is to increase the level of situational awareness a user experiences by first identifying an event of interest (view level 1), then the data sources available (view level 2), then the data

itself (view level 3) [10, 21], making a system that is more suited to analysts trying to understand an entire environment, instead of just individual feeds.

C. Interaction and TUI

Touch functionality is supported with a Kinect as a *Natural User Interface*, enabling multiple users to use the tabletop system simultaneously. The Kinect is mounted approximately 1.5m above the tabletop, with touch events generated from the depth information of the Kinect RGBD images. In addition to touch, users utilize a unique TUI as an interface to control camera feeds within the system. The use of TUIs along with touch enables intuitive control of the system without any specific training.

The TUI used in the SONA system was designed in the shape of a camera as shown in (shown in Fig 1c). The TUI is equipped with several retro-reflective markers to provide 6DOF tracking with an OptiTrack system. The TUI is synchronized with the virtual camera on the map, controlling the panning, tilting, and zoom of the actual CCTV cameras. Users can control the CCTV cameras without any specific training, given the direct mapping between the TUI and the camera. The camera selection is controlled by placing the TUI above a virtual camera's icon or feed on the map. The camera TUI is elevated to avoid obscuring the feed, and given the viewing angle, users are able to view the feed, whilst potentially associating multiple cameras with different feeds.

The TUI camera has five functions: selecting, panning, tilting, zooming, and locking (synchronizing the TUI with a camera). First, selecting is represented as a red circle in Fig. 4e. When the TUI is tracked by the OptiTrack system, a red circle appears under the TUI on the display and symbolizes

the location of the TUI relative to the digital map. In addition, a second marker appears on the corner of the camera feed to indicate it is selected.

- When the TUI has been placed on top of a digital camera marker, the user can rotate the TUI a full 360° (180° left/right depending on the pan abilities of the camera), rotating the selected cameras view to the same absolute direction relative to the digital map (a).
- The TUI can also be tilted, rotating 45° above and below parallel ground view (Fig. 4b).
- Zooming using the TUI is shown in Fig. 4c, and is controlled by rotating the ‘lens’ of the TUI by rotating the three reflective markers around the TUI camera bodies’ central axis. Zooming in is 45° to the right, and zooming out 45° to the left.
- As the camera is moved around the tabletop (Fig. 4d), it becomes associated/disassociated with different cameras on top of which it is placed. Thus, the TUI provides five degrees of freedom (zoom, pan, tilt, x, y).
- When the TUI is within 10cm vertically of a virtual camera on the display, the TUI selects that camera (Fig. 4e). This allows the tangible to be moved either on the table or lifted above and still control the camera up to 10cm before becoming disassociated. The 10cm limit was selected in testing, as we found when changing cameras, users began rotating the camera to the desired position of the new camera whilst the previous camera was still selected. By adding in the vertical limit, we found this greatly reduced the accidental rotations that occurred when moving the TUI’s selection between cameras.

IV. EVALUATION

A user study was conducted comparing the use of SONA’s TUI interaction against the multi-screen system (MS) approach used in existing CCTV centers with traditional inputs. To evaluate the system, we required multiple data sets for environments of varying complexity. Our evaluation used a dynamic virtual city, with multiple virtual cameras, with buildings for obstruction, along with road traffic, and pedestrians. Two virtual worlds were created for the user study; an artificial city (layout and buildings) and a digital recreation of part of a real city (layout and buildings). Further information pertaining to the virtual cities is provided later in this section.

The hypotheses for the study were as follows:

- H1: Participants can maintain focus on an entity of interest for longer periods of time using SONA.
- H2: Participants are required to make less adjustments to the cameras’ pan/tilt parameters using SONA.
- H3: Participants prefer SONA over existing methods.

A. Method

1. Task

Participants performed a surveillance task and were required to monitor and track a vehicle of interest across multiple cameras as it moved around the virtual worlds. The focus of the study was on evaluating the effectiveness of the camera control systems, rather than the direct evaluation of the three-layered/filtered visualization. To ensure the focus

remained on the manipulation of the data, the system was locked to the third level of the visualization, showing all the data streams simultaneously.

The task given to all participants was to “find a suspicious truck and maximize the amount of time it is visible to the cameras by controlling them.” A suspected vehicle (Fig. 7d) was the target and was presented to participants as a printed image showing the truck from two angles. Participants were told to maximize the amount of on-screen time for the truck and on as many cameras as possible.

To compare how effective a participant was at tracking the vehicle of interest, we compared each participant’s tracking time against the maximum possible amount of on-screen tracking time (OST) they could have achieved. Given the different paths of the vehicle for each trial, we converted the OST for each participant into the maximal possible OST (as a percentage) by dividing by the maximum possible OST from the system-controlled cameras. The maximum possible OST was calculated by running artificially-controlled cameras that always focused on the location of the vehicle for the participant’s path. However, OST was only counted when the camera could view the vehicle unobstructed by obstacles such as buildings.

2. Virtual cities

The data feeds for the user study were composed of two 3D environments created in Unity3D: a virtual city (VC) (Fig. 5 left) and a recreation of a real city (RC) (Fig. 5 right) city of Adelaide. The virtual city’s layout consists of a square with two roads intersecting in the middle. This virtual city was composed for the training session and to perform the first trial. The recreation of part of city of Adelaide consisted of a relatively complex structure, being the same layout and utilizing 3D reconstructions for buildings as the real city. As such, RC had six T-junctions and two X-junctions, compared to the VC with four T-junctions and one X-junction. Virtual CCTV cameras were emulated (locations shown in Fig. 6), passing their data feeds to SONA as if from real security cameras, as well as exposing their controls for pan and tilt. Zoom was disabled for the study, as preliminary trials found it unnecessary for the task.

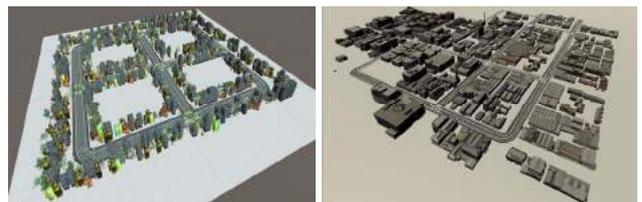


Fig. 5. The virtual city (VC) used for the study (left). The recreation of the real city (RC) used in the study (right).

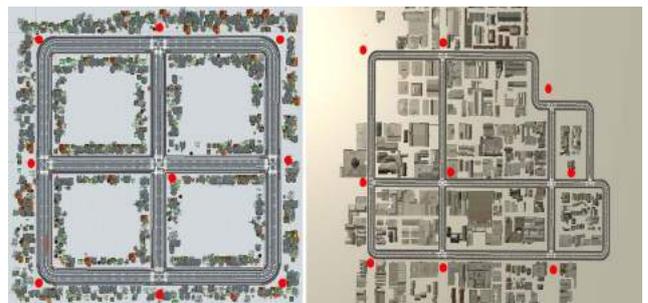


Fig. 6. Location of cameras for the study in VC (left) and RC (right).

3. Apparatus

We used a 47inch monitor for the multi-screen system's display (Fig. 7), and this system displayed all video feed simultaneously. Fig. 7 shows multiple CCTV screens, paper maps, and a suspicious truck, that are all used in the multi-screen condition. The SONA system was implemented as described previously. Both systems used the same computer (Intel Core i7 59030K CPU and nVidia 980TI GPU). For the multi-screen condition, a Logitech Extreme TM X3D Pro joystick was used for camera control, and a Targus numeric keypad was used for camera switching (one camera associated with the numbers one to nine). Six OptiTrack cameras were used to track a single camera TUI.



Fig. 7. (a) The Multi-screen system imitates the existing CCTV center. (b) The virtual city and (c) The recreation of the real city in the study. (d) A suspicious truck used in the study for the task, with traditional paper maps provided to the user.

4. Procedure

To complete all the tasks in the study, participants carried out four steps: preparation, training, the actual study tasks, and post-study questionnaires. In the preparation phase, participants were provided with an information sheet, and if they chose to continue, signed a consent form. After this initial stage, the facilitator explained the tasks to be performed on both systems.

Training on each system was conducted for five minutes (with a five-minute break in between conditions). The presentation of the experimental conditions to participants was counter-balanced. Training for both SONA and MS systems used the VC map. The VC and RC maps comprise of six different vehicle paths in total (four for VC and two for RC). Participants were required to take a five-minute break before proceeding onto the evaluation task.

Following the training session, participants proceeded to the third phase and were evaluated performing the actual task. Both maps (VC and RC) were used in the evaluation trials. The ordering of the experimental conditions was counterbalanced; however, the two tasks were undertaken on the simpler VC map before the more complex RC map. The trials were conducted for ten minutes each with a five-minute break in between trials. The suspicious truck in the trials followed six predefined paths that were randomly generated. Including the training, six trials were run in total (four on the VC, two on the RC) for each participant. Each participant completed all six paths, alternating between systems, and randomized in presentation order.

To evaluate the effectiveness of the geospatial system, we utilized Taylor's *Situational Awareness Rating Technique* (SART) [22], a validated method of measuring situational awareness (SA). The default form of SART concentrates on an understanding of the situation and feeling. The final phase of the study required participants to complete three questionnaires. The first questionnaire (the SART), focused on awareness across both MS and SONA conditions. The questionnaire evaluated the stability, complexity of information, concentration, and familiarity of the environment. The second questionnaire sought system preference information across nine criteria. The last questionnaire asked for feedback about various parts of SONA, along with a section for any other feedback the participants offered. Participants took approximately one hour to complete the entire study.

5. Data collection

Table I describes the data recorded from the surveillance task.

TABLE I. DATA RECORDED FOR EACH TASK.

OST	On screen time	The amount of time the vehicle of interest was visible to each camera. This value could then be compared with the optimal OST for the same camera.
CS	Camera Switch	How many times the user switched the active camera using the UI.
Pan	Panning angle	The total value of the pan control of the cameras.
Tilt	Tilting angle	The total value of the tilt control of the cameras.
P+T	Pan + Tilt	Total value of Pan + Tilt.

6. Participants

Nineteen participants were recruited for the user study from the Ph.D. student body of the University of South Australia or via social networking. The mean age of the participants was 32 (23-42) years old (SD 5.76). The students were familiar with at least basic computer use, with the majority of students having learned three-dimensional expression and visualization techniques in computer science. All participants had no previous experience with CCTV or surveillance systems. They consisted of two females and seventeen males, three participants were left-handed, and sixteen were right-handed. In the 3D gaming experience, 12 students were familiar with 3D gaming environments, and the remaining seven were less experienced. Participants received chocolates as a gift for participating in the study.

B. Results

The results are divided into four sections: surveillance task results, SART questionnaire results, system preference results, and participant feedback.

1. Surveillance Task Results

This section describes the quantitative results of the experiment. Unless otherwise stated, Analysis of Variance (ANOVA) models were used to determine significance. Statistical significance is set at $\alpha = 0.05$. OST as a percentage is calculated by dividing a participant's own OST by the maximum OST possible for that scenario.

On-screen Time: The mean total on-screen time was 64.75% (SD 14.78). Analysis found a significant main effect of system for on-screen time ($F_{(1,18)} = 29.678, p < 0.001$). The mean total on-screen time for the SONA system was 69.52% (SD 15.36), significantly more than the mean total time for the multi-screen system 59.97% (SD 12.64). Participants tracked the suspicious object for longer using SONA than MS. There was a significant main effect of the type of map used on on-screen time ($F_{(1,18)} = 29.398, p < 0.001$). The mean total OST while using VC was 59.76% (SD 13.57), which was significantly less than RC 69.73% (SD 14.41). A significant interaction effect between the system type and map type was not found.

Camera Switch: Total mean camera switch (CS) count was 56.63 (SD 23.12). System type had a significant main effect on the camera switch count ($F_{(1,18)} = 9.593, p < 0.01$). SONA's mean total camera switch count was 61.36 (SD 24.59), significantly higher than MS with 51.87 (SD 20.78). However, map type and interaction between system and map used reported no significant findings.

Pan: The mean total pan was 23883.43 degrees (SD 22523.60). Analysis found a significant main effect of system on pan ($F_{(1,18)} = 22.181, p < 0.001$). The mean total of the pan movement for the multi-screen system was 12563.70 degrees (SD 3595.69), significantly different from that of SONA reporting 35203.16 degrees (SD 27426.42). The map type had no significant effect on the amount of panning, and no significant interaction between system and map was found.

Tilt: The mean total tilt was 6384.89 degrees (SD 9642.37). Analysis revealed no significant findings.

Pan + Tilt: The mean total pan and tilt (P + T) was 30268.32 degrees (SD 28231.80). There was a significant main effect of the system used on pan and tilt ($F_{(1,18)} = 13.323, p < 0.01$). The mean total for pan and tilt for the multi-screen system was 18768.68 degrees (SD 7535.90), significantly different from that of SONA reporting 41767.96 degrees (SD 35877.42). No other significance was found.

2. SART Questionnaire Results

Participants were asked to describe their experience with MS and SONA systems via an exit questionnaire. Participants were also asked to rate the systems on a seven-point scale in terms of interaction with the systems. Results from the SART questionnaire [22, 23] are as follows. Unless otherwise stated, Wilcoxon Signed-rank tests were used to determine significance. Statistical significance is set at $\alpha = 0.05$.

Q1: Instability of Situation: “*Is the situation highly unstable and likely to change suddenly (High) or is it very stable and straight forward (Low)?*” is not significantly different between SONA and MS. Lower scores indicated ease of use.

Q2: Complexity of Situation: “*Is it complex with many interrelated components (High) or is it simple and straight forward (Low)?*” is as follows: participants found that situations presented in SONA (Mean = 2.47, Mdn = 2.00) were significantly more simple than environments viewed using MS (Mean = 3.26, Mdn = 3.00, $z = -2.30, p = 0.02, r = -0.37$). Lower scores indicated ease of use.

Q3: Variability of Situation: “*Are there a large number of factors varying (High) or are there very few variables changing (Low)?*” is as follows: participants found that SONA (Mean = 2.79, Mdn = 3.00) was significantly less complex than the multi-screen system (Mean = 3.63, Mdn = 3.00, $z = -2.22, p = 0.03, r = -0.36$). Lower scores indicate ease of use.

Q4: Arousal: “*Are you alert and ready for activity (High) or do you have a low degree of alertness (Low)?*” is as follows: participants were significantly more alert in environments with SONA (Mean = 4.84, Mdn = 6.00) than the MS system (Mean = 3.89, Mdn = 4.00, $z = -2.54, p = 0.01, r = -0.41$). Higher scores indicated better alertness.

Q5: Concentration of Attention: “*How much are you concentrating on the situation? Are you concentrating on many aspects of the situation (High) or focused on only one (Low)?*” is as follows: the use of SONA (Mean = 5.16, Mdn = 6.00) significantly allowed participants to concentrate on many more aspects of the situation system than with the multi-screen system (Mean = 4.05, Mdn = 5.00, $z = -2.06, p = 0.04, r = -0.33$). Higher scores indicated heightened situation awareness.

Q6: Division of Attention: “*How much is your attention divided in the situation? Are you concentrating on many aspects of the situation (High) or focused on only one (Low)?*” is not significantly different between SONA and MS. Higher scores were better.

Q7: Spare Mental Capacity, “*How much mental capacity do you have to spare in the situation? Do you have sufficient to attend to many variables (High) or nothing to spare at all (Low)?*” is not significantly different between SONA and MS. Higher scores indicated heightened situation awareness.

Q8: Information Quantity: “*How much information have you gained about the situation? Have you received and understood a great deal of knowledge (High) or very little (Low)?*” is as follows: participants found that SONA (Mean = 5.16, Mdn = 5.00) significantly provided them with greater information pertaining to the situation than the multi-screen system (Mean = 4.58, Mdn = 5.00, $z = -2.65, p = 0.01, r = -0.43$). Higher scores indicated heightened situation awareness.

Q9: Familiarity with Situation, “*How familiar are you with the situation? Do you have a great deal of relevant experience (High) or is it a new situation (Low)?*” is not significantly different between SONA and MS. Lower scores indicated heightened situation awareness.

3. System Preference Results

Participants were asked to rank the systems for each of the following questions:

- Q1. Which system is easier to understand the direction of the target for each task?
- Q2. Which is faster get the target on the screen?
- Q3. What was more enjoyable (fun/interesting/happy) during the task performed?
- Q4. Which system identified the next camera of the target as you expected?

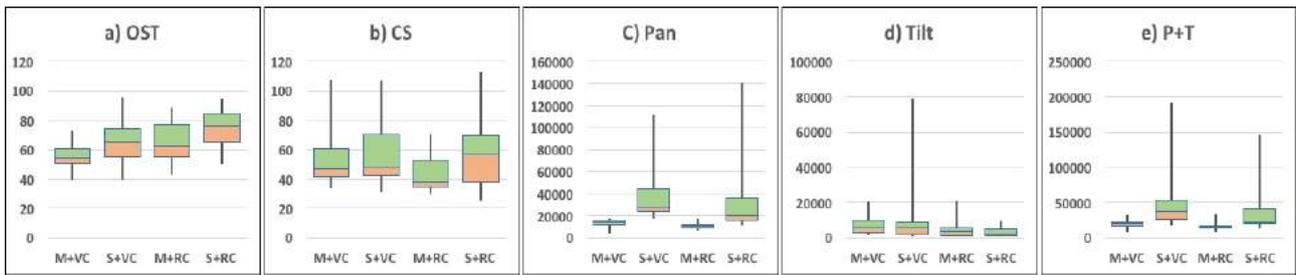


Fig. 8. Quantitative results of the surveillance task showing: a) OST, b) CS, c) Pan, d) Tilt, and e) Pan+Tilt.

- Q5. Regarding panning/tilting which system more effort to find the target?
- Q6. Which did you understand more Joystick or Tangible User Interface (TUI)?
- Q7. Which User Interface is more intuitive to control?
- Q8. Overall which visualization did you prefer to use?

Fig. 9 represents preference counts for using each system. For example, for question 1, 15 participants ranked SONA as being easier to understand for tracking the direction of the target versus 4 participants who ranked the multi-screen system as being easier. For every question, participants preferred SONA over MS. However, in questions 5, 6, and 7, MS and SONA marked a similar score. Presumably due to users being already familiar with Joystick and Keypad.

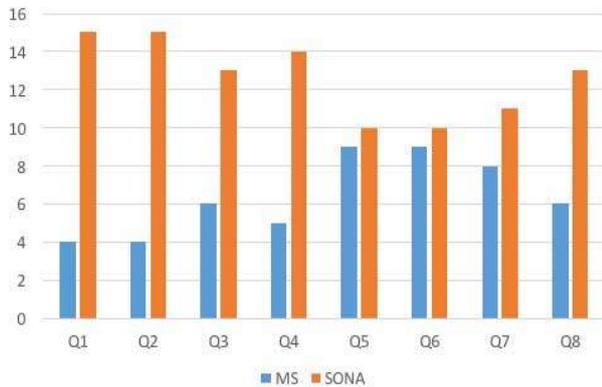


Fig. 9. System preference

4. Participant feedback

In addition to the two questionnaires, we included four questions for written feedback. The first question asked the participants' opinion regarding what SONA did well in the perspective of functions. The second question required participants to gauge the best aspects of SONA from a conceptual perspective. The third question asked participants to comment on the limitations of SONA. The fourth question was an open question for any other feedback. Notable responses are summarized as follows. When asked about what SONA did well, participants responded that the view made it easy to use, whilst the use of the TUI with the visualization provided a good combination. Moreover, one participant said they felt using SONA was like playing a computer game. In answering the second question, participants stated

that SONA was easy and fast to control. A limitation described by participants was that their hand and TUI controller could sometimes obscure the screen; however, this problem would also be present in other display/TUI systems. With regards to the general feedback received, participants suggested putting the CCTV screen above the TUI on the LCD display to avoid obscuring it and using multiple TUIs.

C. Discussion

As demonstrated by the quantitative (i.e. Pan, Tilt, P + T) and qualitative results (1, 2, 6, 7, 8 of the preference system survey), participants maintained a longer OST tracking the suspected vehicle using SONA (supporting H1). Although users took the same amount of time for each task, SONA was more effective in tracking the vehicle. MS and SONA have similar difficulty levels as reported in the system preference survey. Participants found SONA interactions more intuitive. This shows that SONA is not only superior in terms of results, but also in terms of user preference (supporting H3).

Participants performed with a statistically higher number of interactions using SONA when compared to MS (not supporting H2). Despite having a higher number of interactions, SONA resulted in a higher amount of OST. One explanation for this might be that users felt they could make continual, minor adjustments to optimize tracking, instead of less regular, gross movements.

Situational awareness measured by the SART questionnaire [24], found SONA to be superior to MS for five of the nine criteria, with the remaining four not showing any significant difference. These results again favor SONA's approach to interacting with the data. The responses indicate that participants found it easy to concentrate on and to understand the situation. Two possible reasons participants preferred the SONA approach to data examination are the straightforward and precise operation of the TUI and the synchronized operation of the angles of the actual cameras, which provide an intuitive interface for interactions.

Several technical limitations with SONA were identified from the user study. One piece of feedback surrounded using multiple TUIs, however, given the use of optical tracking markers, our previous attempts at getting multiple TUIs working in shared space was problematic. This is a limitation of the tracking technology used and not the SONA approach.

V. CONCLUSION

We introduced SONA, a novel method for exploring geotagged data. We presented our implementation within the context of surveillance and CCTV exploration. We developed to explore a virtual city with virtual CCTV cameras using a three layered geo-visualization, with abstract events represented (e.g., car accidents), the location of data sources near that event, and then the final data streams, shown for each of the three layers. These cameras are accessed through a TUI.

To evaluate the effectiveness, we compared SONA against the current multiscreen approach for navigating CCTV data. The results of the study showed the SONA's methods for interacting with geotagged data are superior to existing methods. In addition to providing a superior method for interacting with the data, SONA provided increased situation awareness and understanding for users.

As a future work, the main direction is exploring SONA in a collaborative setting. In particular, we will focus on the development of tools to help professionals in a different domain (with different forms of streaming data), focusing on other applications such as exploring social media. We will explore the application of active TUIs, enabling the synchronization of input controls between remote users using the same system.

ACKNOWLEDGMENT

This research has been supported by the D2D CRC (Data to Decisions Cooperative Research Centre) whose activities are funded by the Australia Commonwealth Government's Research Centre Programme.

REFERENCES

- [1] X. Wang, P. E. Love, M. J. Kim, and W. Wang, "Mutual awareness in collaborative design: An Augmented Reality integrated telepresence system," *Computers in Industry*, vol. 65, no. 2, pp. 314-324, 2014.
- [2] V. Maquil, "Towards Understanding the Design Space of Tangible User Interfaces for Collaborative Urban Planning," *Interacting with Computers*, p. iwv005, 2015.
- [3] Y. Qin, J. Liu, C. Wu, and Y. Shi, "uEmergency: a collaborative system for emergency management on very large tabletop," presented at the ACM ITS, Cambridge, Massachusetts, USA, 2012.
- [4] S. Doeweling, T. Tahiri, P. Sowinski, B. Schmidt, and M. Khalilbeigi, "Support for collaborative situation analysis and planning in crisis management teams using interactive tabletops," in *Proc. ACM ITS*, 2013, pp. 273-282.
- [5] W. W. Gaver, "Technology affordances," in *Proc. SIGCHI*, 1991, pp. 79-84.
- [6] M. Taha, "Parramatta Park stabbing: CCTV roll-out may include place where Prabha Arun Kumar was killed," <http://www.abc.net.au/news/2015-03-25/cctv-boost-could-include-parramatta-park-prabha-arun-kumar/6347732>, 2015.
- [7] A. Cunningham, B. Close, B. Thomas, and P. Hutterer, "Design and impressions of a multi-user tabletop interaction device," presented at the Proceedings of the Eleventh Australasian Conference on User Interface - Volume 106, Brisbane, Australia, 2010.
- [8] A. Anwar, T. Nagel, and C. Ratti, "Traffic origins: a simple visualization technique to support traffic incident analysis," in *PacificVis* 2014, pp. 316-319.
- [9] C. Fang *et al.*, "ViCAT: visualisation and interaction on a collaborative access table," in *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems (T-ABLETOP '06)*, 2006.
- [10] B. Shneiderman, "The eyes have it: A task by data type taxonomy for information visualizations," in *Proc. IEEE Symposium*, 1996, pp. 336-343.
- [11] D. P. Käser, M. Agrawala, and M. Pauly, "FingerGlass: efficient multiscale interaction on multitouch screens," in *Proc. SIGCHI*, 2011, pp. 1601-1610.
- [12] M. Masoodian, S. Luz, and D. Kavenga, "nu-view: a visualization system for collaborative co-located analysis of geospatial disease data," in *Proc. ACSW*, 2016, p. 48.
- [13] C. Guerlain, S. Cortina, and S. Renault, "Towards a collaborative Geographical Information System to support collective decision making for urban logistics initiative," *Transportation Research Procedia*, vol. 12, pp. 634-643, 2016.
- [14] Y. Ming, J. Jiang, and F. Bian, "3D-City Model supporting for CCTV monitoring system," *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, vol. 34, no. 4, pp. 456-459, 2002.
- [15] S. Alaçam, Y. İpek, Ö. Balaban, and C. Kayalar, "Organising crowd-sourced media content via a tangible desktop application," in *MultiMedia Modeling*, 2014, pp. 1-10.
- [16] G. W. Fitzmaurice, H. Ishii, and W. A. Buxton, "Bricks: laying the foundations for graspable user interfaces," in *Proc. SIGCHI*, 1995, pp. 442-449: ACM Press/Addison-Wesley Publishing Co.
- [17] G. W. Fitzmaurice and W. Buxton, "An empirical evaluation of graspable user interfaces: towards specialized, space-multiplexed input," in *Proc. SIGCHI* 1997, pp. 43-50.
- [18] J. A. Walsh, S. v. Itzstein, and B. H. Thomas, "Tangible agile mapping: ad-hoc tangible user interaction definition," presented at the Proceedings of the Fourteenth Australasian User Interface Conference - Volume 139, Melbourne, Australia, 2013.
- [19] J. Mateu, M. J. Lasala, and X. Alamán, "Tangible Interfaces and Virtual Worlds: A New Environment for Inclusive Education," in *Ubiquitous Computing and Ambient Intelligence. Context-Awareness and Context-Driven Interaction*: Springer, 2013, pp. 119-126.
- [20] C. E. Jones and V. Maquil, "Towards Geospatial Tangible User Interfaces: An Observational User Study Exploring Geospatial Interactions of the Novice," *GISTAM*, vol. 582, pp. 104-123, 2016.
- [21] I. Herman, G. Melançon, and M. S. Marshall, "Graph visualization and navigation in information visualization: A survey," *IEEE TVCG*, vol. 6, no. 1, pp. 24-43, 2000.
- [22] R. Taylor, "Situational Awareness Rating Technique(SART) : The development of a tool for aircrew systems design," *A GARD, Situational Awareness in Aerospace Operations 17p(SEE N 90-28972 23-53)*, 1990.
- [23] P. M. Salmon *et al.*, "Measuring Situation Awareness in complex systems: Comparison of measures study," *International Journal of Industrial Ergonomics*, vol. 39, no. 3, pp. 490-500, 2009.
- [24] M. R. Endsley, "Situation awareness global assessment technique (SAGAT)," in *Proc. IEEE, NAECON*, 1988, pp. 789-795.