User Study on 3D Multitouch interaction (3DMi) and Gaze on Surface Computing

[1]Eugene Ch'ng, [2]Neil Cooke

[1]School of Computer Science, International Doctoral Innovation Centre, University of Nottingham Ningbo China, 199 Taikang East Road, Zhejiang, Ningbo 315100 China
[2]Department of Electronic, Electrical and Computer Engineering, University of Birmingham, Edgbaston, B15 2TT, United Kingdom
eugene.chng@nottingham.edu.cn, n.j.cooke@bham.ac.uk

Abstract. On a multitouch table, user's interactions with 3D virtual representations of real objects should be influenced by task and their perceived physical characteristics**.** This article explores the development and user study of an interactive 3D application that allows users to explore virtual heritage objects on a surface device. To-date, most multitouch has focused on 2D or 2.5D systems. A user-study is reported where we analyse their multimodal behaviour – specifically how they interact on a surface device with objects that have similar properties to physical versions and the users' associated gaze patterns with touch. The study reveals that gaze characteristics are different according to their interaction intention in terms of position and duration of visual attention. We discovered that virtual objects afford the perception of haptic attributes ascribed to their equivalent physical objects, and that differences in the summary statistics of gaze showed consistent characteristics between people and differences between natural and task based activities. An awareness of user behaviours using natural gestures can inform the design of interactive 3D applications which complements the user's model of past experience with physical objects and with GUI interaction.

Keywords: Interactive 3D, multitouch, surface computing, digital heritage, gaze tracking

1 Introduction

Multitouch surface computing in public spaces dedicated to heritage such as museums provide the opportunity to enhance people's experience, affording social interaction with others around digitised knowledge sources and virtual artefacts. The broad goal of this research is to recreate virtual experience of heritage objects, with the purpose of drawing 3D digital heritage objects from the archives for public access. This paper addresses an important sub-goal – to understand how users would behave when given 3D objects to manipulate within the virtual environment, through multi-touch gestures on a surface computer in the social space of a museum. Specifically,

how users would manipulate 3D objects that have simulated properties similar to their physical versions (physics effects, collisions, weight, etc.) given simple tasks. In addition to capturing gestures, we measure the user's gaze direction with an eye tracker in order to better understand how a person's visual attention is allocated during multitouch gesture. Thus, we aim to show that on a multitouch table, users' interactions with 3D virtual representations of real objects should be influenced by task and their perceived physical characteristics.

The article begins with a background of this particular research and its interest. The motivation for access to heritage artefacts, particularly those from the archives is discussed. The next section reviews related topics. The article continues with the methodology, and the results and discussions section, which describe the subject and development of the multitouch application with observations from user-evaluation. The article ends with a conclusion and future direction.

2 The Virtual Within the Physical Space

Surface computing with simultaneous multitouch inputs adds a new dimension to the access of information. Initially, applications using surface computing were used for browsing images and videos. These were applications with very basic functionalities (see examples [5, 6]). However, the new paradigm needs new explorations in user interface design that incorporates collaborative features and user evaluation. As surface hardware, APIs and SDKs mature, more creative use will be expected.

One of the critical ways in which digital heritage objects can be made more accessible to a wider audience is the development of more intuitive user interfaces. The touch and gesture-based smartphones and tablet computers have, to date, 'taught' massive amounts of users the multitouch, gesture-based interaction model. It has revolutionised the way in which users access information. Larger touch-screens such as the iPad are allowing a wider set of gestures, e.g., the navigation between Apps of up to 4 fingers using the 'swipe' gesture as opposed to the PC-era 'Alt-TAB' key combinations on the keyboard. These developments are revolutionising both work and leisure. Computers are now intuitive for a broad range of potentially cyberphobic audiences that never knew the PC-era. Computers are for the first time useful and fun, as evident in news channels and magazines that interviewed the elderly of their experience of such devices.

The commercialisation of horizontally oriented tabletop computers such as Microsoft's Surface, PQLab, and Ideum's multitouch, multiuser (MTMU) tabletop computers for museum spaces are bringing general and research computing into another dimension. Large High Definition (HD) displays of up to 65" supporting up to 32 touches and 'pop-out' 3D stereographics already exists (the Digital Humanities Hub-commissioned Mechdyne MTMU tabletop computer at the Chowen Prototyping Hall, the University of Birmingham). The fusion of cutting-edge technological advancements on tabletops are ushering in functional capabilities that were not present in traditional computing environments. Traditional computing environments are sequential, with supposedly 'collaborative' tasks passed between workers either via email or via a single-display, single input terminal. Although concurrent

versioning systems and computer supported collaborative work are available [8], there were issues [7] associated with it, particularly via a single user terminal. Working together on location might be better as it resolves issues of psychological ownership and perceived document quality, as evident in a collaborative Google Doc study [1]. Multitouch and multiuser surface computing opens up possibilities where collaboration is transformed from sequential to simultaneous – all workers work on a task simultaneously. In this sense, learning and access of information becomes more natural.

Research have shown that direct-touch interfaces do evoke confusion for first time users using touch interaction, organisation of content, and occlusion in uncontrolled environments [11]. More recent research suggests that surface computing are providing scopes for interactions that are nearer in experience to physical interactions as compared to classical windowed interfaces [9]. Will users resort to physical interaction models on surface computing? The answer is no, on both past studies and our observations in the present research. Users are influenced by the desktop paradigm. Research on user-defined gestures in surface computing [12] suggests that the Windows desktop paradigm has a strong influence on users' mental models; that users rarely care about the number of fingers they employ; that one hand is preferred to two, and that on-screen widgets are needed. In our observation of user evaluation conducted in past open days and the present research, users are also influenced by the touch-based smartphones and tablet paradigms.

The behaviour of large crowds in uncontrolled environment suggests that users learn from each other. An observation [10] with 1199 participants reveals that users at a display attract other users, and a user's actions on the touch wall is learned by observers. An interesting result was "how these people were configured in groups of users and crowds of spectators rather than as individual users. They were able to use the display both in parallel and collectively by adopting different roles" – the use of the display was highly non-individualistic.

Whilst single and multiple user interactions have been studied to a certain extent, 3D Multitouch interaction (newly abbreviated here as 3DMi) is an entirely new area that is yet to be fully explored. 3D interaction in multitouch was briefly mentioned in 2008 by Bowman *et. al.* [2], "The current trend towards multi-touch interfaces at least acknowledges that humans tend to act with more than one finger at a time, but still this is just scratching the surface of the immersive experience that virtual environments will offer in future computer applications. What about grasping, turning, pushing, throwing, and jumping when interacting with computer applications?" Indeed, intuitive 3DMi has a long way to go, but there needs to be a new initiative for research here considering that market trends have changed since 2008 with more demands for multitouch surface computing worldwide.

3   Methods

A surface computing 3DMi application was developed. The 3DMi incorporates 3D objects that simulate weight, friction and gravity. More details on the implementation can be acquired from two articles [3, 4].

To identify and distinguish gesture behaviours, 9 participants (A to I) were monitored while they interacted with 3DMi in distinct phases whilst wearing the Tobii Eye Glasses for capturing monocular gaze position (point of view 56° horizontal and 40° vertical). 30 infrared markers were placed equidistant around the edges of the table. A separate video camera records the interactions. Gaze data for each mode and participant were analysed:

1. *Passive Gaze Observation*: participant listens and watches the Instructor.
2. *Active (Free Exploration)*: participant is free to explore and manipulate the objects on the table with no explicit aim.
3. *Active (Task-Specifics)*: the participant is given a specific task requiring the manipulation of the artefacts on the table to fulfil an educational objective.

The sections below present our findings.

3.1 Observations

Virtual objects do simulate the perceived haptic attributes of real objects (weight, surface textures). Due to the realistic physics simulation, observation of user interaction with objects suggests that their perception of the digital facsimiles correlated with that of physical objects:

- Dexterity is observed where quick learners (D) picked up gestures where tasks are accomplished quickly through taking advantage of the weight, size and the effects of gravity and velocity of the object – flicking objects to the intended location.
- The larger the virtual object, the less likely it will be pushed aside (A, C).
- Participants (E) pushed obstacles aside with the other hand whilst moving the task object to the destination.
- A correlation between the number of fingers used and the perceived weight of the objects (B, D).
- When there is friction (object resists movement), participants pressed down more heavily on the surface
- Double tap objects to select, a behaviour learned from mouse use. (D, F).
- Exploration of gesture limits. For example the extent of the zoom, the speed at which objects can be dragged (All).
- While moving virtual objects, users pass objects from one hand to another (All).

The following gaze behaviours were observed on all participants:

- Gaze follows an object when dragged; gaze is depended upon as there were no haptics on the touch screen.
- Head is oriented so that focus of touch is in the centre of vision (central bias).
- Gaze is rarely focused upon the hand but on the visible part of the underlying object.

- If both hands are dragging objects in the same direction, gaze will tend to fall on the object nearest to the target. If objects are dragged to different targets, then gaze will fall between them or onto their point of convergence (Fig. 1)
- Gaze is a reliable predictor of where the person will touch next (i.e. the next object to be grabbed) (Fig. 2)
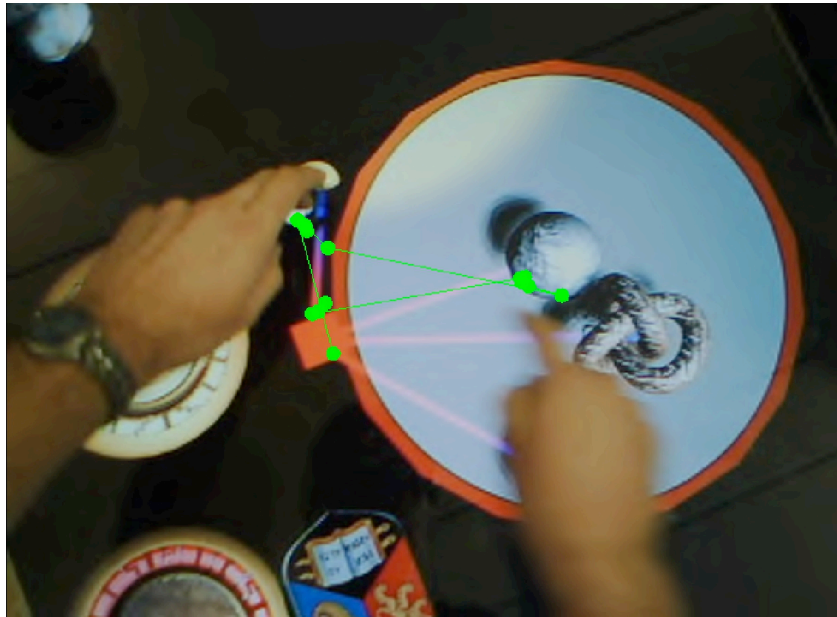

Figure 1. A participant's gaze patterns (in green) over a 0.5 window while conducting a multitouch gesture. Gaze moves between the two objects and their origin (the red square).
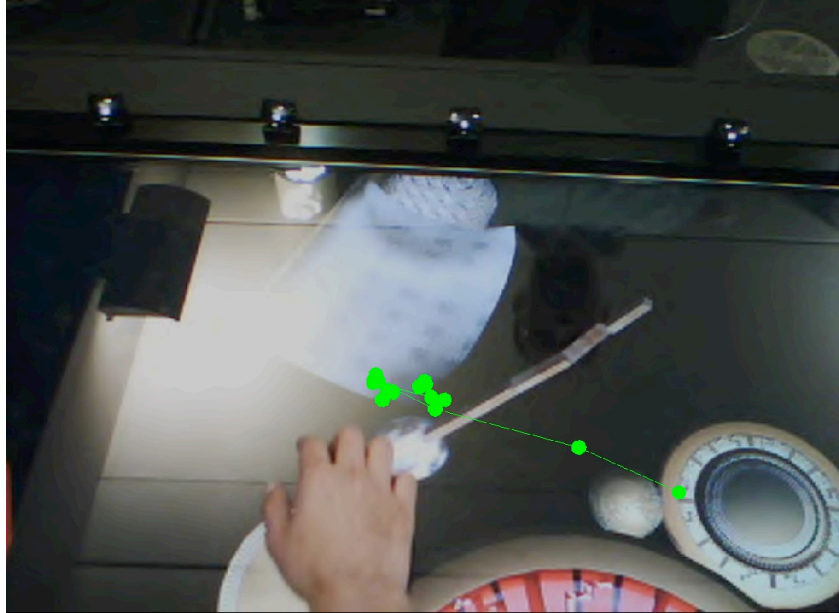
Figure 2. Gaze tends to follow the object being dragged with fixations towards next object to be touched (in this example the large disc on the right).

3.2 Gaze Characteristics

Gaze characteristics are different in interaction modes for visual attention, position and duration (Table 1). Overall, passive interaction resulted in the shortest fixations (mean – 0.41s, stdev = 0.35s N=640). Longer fixation durations are observed when the participants are actively using the table (mean = 0.64s, stdev=0.88s, N=715), with the imposition of a definitive task shortening the mean duration and its variance (mean=0.52s, stdev=0.71s, N=820).

Fixation positions also showed a difference (Figure 3). For active interaction (free and passive), visual attention is focused on a position between the hands, particularly when the interaction is free. In passive mode, visual attention has a wider spread. Taken together, differences in gaze can be attributable to task.

Differences in the summary statistics for gaze showed consistent characteristics between people and differences between natural and task based activities. This suggests that the natural state of interaction the application affords (free play state) and specific task based interaction states that can be inferred from gaze alone. These are preliminary results. Gaze characteristics can thus potentially be used in inference models to deduce the tasks undertaken by museum visitors and predict touch gestures allowing for applications to prime relevant information for access.

	Passive			Active - Free			Active - Task		
Participant	Sample Count	Mean	Standard Deviation	Sample Count	Mean	Standard Deviation	Sample Count	Mean	Standard Deviation
A	140	0.38	0.31	166	0.67	0.84	168	0.45	0.34
B	77	0.39	0.22	58	1.1	1.91	94	0.84	1.46
C	-	-	-	86	0.66	0.63	94	0.44	0.32
D	89	0.49	0.4	90	0.54	0.81	126	0.46	0.39
E	121	0.41	0.34	104	0.4	0.34	131	0.52	0.69
G	91	0.48	0.54	118	0.76	0.88	93	0.66	0.9
I	122	0.34	0.24	93	0.48	0.37	114	0.38	0.28
All	640	0.41	0.35	715	0.64	0.88	820	0.52	0.71

Table 1 Summary gaze statistics showing fixation duration distribution estimations per phase, per participant. All participants exhibited shorter fixation distributions with smaller standard deviation when actively engaged in the task, compared to when freely interacting with objects. The shortest fixation durations occur when users are not gesturing (passive mode). All times are in seconds.
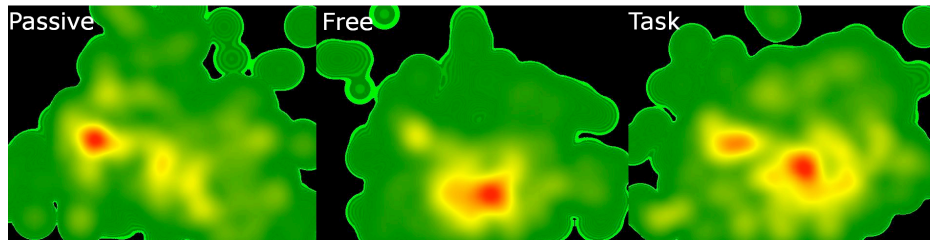

Figure 3. Heat map visualisation of visual attention on touch table from all participants' point of view for the three different phases. Red shows the highest concentration of gaze, green the lowest with black showing no gaze. Active use of multitouch (Free and Task) show a concentration of attention in the middle towards the bottom related to manipulating objects between hands, with differences between Free and Task indicating a more dynamic exploration for Task, as the red is more dispersed. Passive (no multitouch) does not have a central concentration of fixations because users are not using their hands.

5 Conclusion

In this article, we presented our findings on the multimodal behaviour and gaze of users during 3D multitouch interaction, with a broad goal of recreating virtual experience of heritage objects, and a sub-goal of understanding user behaviour when given 3D objects to manipulate. The research has direct relevance to the access of heritage objects via digital means, which have important economic and social value. Heritage contributes directly and indirectly to the GDP of a country that hosts them and the public access and valorisation of heritage promotes the artistic, aesthetic, cognitive and recreation needs for individuals, households, and their national identity. Unrestricted access of heritage from the archives via digital interfaces allows the rediscovery of hidden source of information that may bridge relationship or

chronological gaps amongst artefacts. The introduction of virtual information spaces hosting realistic laser-scanned 3D objects rendered in interactive real-time computer graphics, coupled with natural gestures in 3D multitouch screens are one of the important and accessible ways of interacting with heritage objects. These virtual environments occupy a little space (65" screens mounted vertically, or as table computers) and complement the limitations of space in museums, but the value that they are able to add to the learning, teaching, research, and access of heritage is significant.

In this article, we investigated how multitouch surface computing can contribute to the research and social interaction opportunities of accessing heritage objects to enhance users' experience around digitised knowledge sources and virtual artefacts. We explored the development and user study of a 3DMi application that allows users to explore virtual objects using natural gestures. Our study allows us to analyse their multimodal behaviour – specifically how they interact on a surface computer with objects that have similar properties to their physical versions, and the users' associated gaze patterns with touch.

We showed that on a multitouch table, users' interactions with 3D virtual representations of real objects are influenced by task and their perceived physical characteristics. Gaze characteristics are different according to interaction modes in terms of the allocation of visual attention. Virtual objects can afford haptic attributes of physical objects, although users may revert to old interaction modes from the Windows GUI era suggesting that the perception of affordance by system designers should not be assumed. Differences in the summary statistics for gaze demonstrate consistent characteristics between people, and differences between natural and task based activities. An awareness of how objects afford interaction in a natural state can inform design in order to encourage constructive activities.

Our study is an initial step of a broader goal to understanding user behaviour and multimodal interaction with 3D objects on surface computers. We believe the findings articulated in this research will contribute to better design of 3D multitouch applications using natural gestures.

Future studies will involve a redesign of the interactive 3D application to compensate for users' perception of virtual objects in relation to their understanding of the haptics and physics of real objects. We aim to also conduct studies on multiuser and multitouch collaborative tasks involving two, and up to four users in the evaluation to gain understanding of how users behave in a collaborative digital table, monitoring gaze patterns to assist in resolving gesture intent.

Acknowledgements

References

[1]   Blau, I. and Caspi, A. 2009. What Type of Collaboration Helps? Psychological Ownership, Perceived Learning and Outcome Quality of Collaboration . *Proceedings of the Chais conference on instructional technologies research 2009: Learning in the technological era* (Raanana, 2009), 48–55.

[2]   Bowman, D.A. 1999. Interaction techniques for common tasks in immersive virtual environments. Citeseer.

[3]   Ch'ng, E. 2012. New Ways of Accessing Information Spaces Using 3D Multitouch Tables. *Proceedings of the Art, Design and Virtual Worlds Conference, Cyberworlds 2012, 25-27 September 2012* (Darmstadt, Germany, 2012).

[4]   Ch'ng, E. 2013. The Mirror Between Two Worlds: 3D Surface Computing Interaction for Digital Objects and Environments. *Digital Media and Technologies for Virtual Artistic Spaces*. IGI Global.

[5]   Ciocca, G. et al. 2012. Browsing museum image collections on a multi-touch table. *Information systems*. 37, 2 (2012), 169–182.

[6]   Correia, N. et al. 2010. A multi-touch tabletop for robust multimedia interaction in museums. *ACM International Conference on Interactive Tabletops and Surfaces*. ACM.

[7]   Dekeyser, S. and Watson, R. 2006. *Extending Google Docs to Collaborate on ResearchPapers*. The University of Southern Queensland.

[8]   Eseryel, D. et al. 2002. Review of computer-supported collaborative work systems. *Education Technology & Society*. 5, 2 (2002).

[9]   North, C. et al. 2009. Understanding multi-touch manipulation for surface computing. *Human-Computer Interaction–INTERACT 2009*. Springer. 236–249.

[10]   Peltonen, P. et al. 2008. It's Mine, Don't Touch!: interactions at a large multi-touch display in a city centre. *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*. ACM.

[11]   Ryall, K. et al. 2006. Experiences with and Observations of Direct-Touch Tabletops. *Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems, TABLETOP '06*.

[12]   Wobbrock, J.O. et al. 2009. User-defined gestures for surface computing. *Proceedings of the 27th international conference on Human factors in computing systems*. ACM.