Kolab: appropriation & improvisation in mobile tangible collaborative interaction

Conference or Workshop Item

How to cite:

Dalton, Nicholas; Mackay, Gordon and Holland, Simon (2012). Kolab: appropriation & improvisation in mobile tangible collaborative interaction. In: DIS 2012 Designing Interactive Systems, 11-15 Jun 2012, Newcastle, UK.

Version: Accepted Manuscript

Link(s) to article on publisher's website:
http://dx.doi.org/doi:10.1145/2317956.2317960
http://www.dis2012.org/

Kolab: appropriation & improvisation in mobile tangible collaborative interaction

First Author Name (Blank if Blind Review)
Affiliation (Blank if Blind Review)
Address (Blank if Blind Review)
e-mail address (Blank if Blind Review)
Optional phone number (Blank if Blind Review)

Second Author Name (Blank if Blind Review)
Affiliation (Blank if Blind Review)
Address (Blank if Blind Review)
e-mail address (Blank if Blind Review)
Optional phone number (Blank if Blind Review)

ABSTRACT
Current design guidelines for conventional tangible systems suggest that the *representational significance* of tangible tokens is an important consideration in the design of tangible interaction, especially in collaborative contexts. Such advice might be assumed to imply that nomadic systems that employ improvised tokens are liable to have highly impaired usability. In this paper we describe Kolab, a prototype for a nomadic tangible interaction system that permits any surface to be appropriated as a collaborative tabletop, and which affords the use of a wide range of appropriated artifacts as improvised tangibles. We demonstrate an approach for realizing the necessary interaction techniques combining tangibles and hand gestures using a fusion of image and depth sensing. We present the results of a user study showing that while users' choices of artifacts were seen to follow an unexpected pattern, various artifacts were appropriated and improvised as tangibles, and the system was found to be both usable and well able to support user collaboration.

Author Keywords
Tangibles, mobile, interaction, appropriation, collaboration.

ACM Classification Keywords
H5. User interfaces, H 5.3 collaborative computing

General Terms
Human Factors, Algorithms, Design,

INTRODUCTION
Tangible user interaction (TUI) [4, 5, 13] uses physical artifacts to portray and manipulate digital information. The general aim of the present study is to broaden our understanding of tangible interaction in the context of nomadic, co-located, and cooperative working using shareable interfaces. By 'nomadic' we mean portable systems that are static when operated; laptops are an example of nomadic computing. In real world uses users tend to use more than a single application and so it seems natural to expect future nomadic tangible devices do the same, leaving a potentially useful and relatively unexplored design space. With typical existing approaches to tangible interaction, multiple applications would require multiple tangible tokens sets [1], limiting the portability of such a system. Also, the tangible sets would be prone to attrition by the loss or mislaying of potentially critical tokens [14](section 9.2.1) or simply getting mixed between applications.

Our proposal is that one alternative is to take advantage of the fact that the world is full of potential tangibles and that people appear to be comfortable in improvising with them, as suggested by [2]. Yet the tangible design literature (e.g. [5,6,7,8]) strongly suggests that artifacts need take on '*explicit physical form*' [8] and that the 'representational significance'[7] (the degree of match between abstract data and physical form) of a physical token(or phicon) is a vital part of its representational legibility for users. Fitzmaurice[5] also differentiates between 'weak general' and 'strong specific' tokens to the extent that he defines a 'strong specific' tokens as a basic property of a graspable interface. Holmquist et al [6] differentiate between 'containers' and 'tokens' the latter need 'physically resemble the information they represent in some way' and use tokens as the basis for their 'Token-Based access to digital information frame work'.

It is this strength of opinion that strongly questions the usability and mnemonic resilience of systems that rely on user-improvised tokens for *co-located collaboration*. Thus, the purpose of our study was to consider the question of whether user-selected improvised tokens that have marginal representational significance impair shared tangible interaction to an extent where the viability of nomadic co-located collaboration is compromised.

Previous work
In this section we compare the general interaction framework of Kolab with that of some related systems. (For details of the UI specifically used in the experiment, see next section.) For reasons of space, we will restrict our critical review to broadly nomadic systems. Kolab is designed to be a topdown-projected computer-vision-based interactive table system, and in this respect is comparable to Docklamp [11], PlayAnywhere [15], Visual Touchpad [13], Bonfire [10], reacTIVision [9], FACT [12] and Portico [20]. Like some of these systems, Kolab implements finger tracking both for gestural input and multi-touch interaction, however, Kolab differs by its use of *marker-less* tangibles

DIS 2012, June 11-15, 2012, Newcastle, UK.

by using 3D sensing. Like Visual TouchPad [13], Kolab uses 2.5D depth sensing along with image capture to refine gesture recognition, tangible and multi-touch input but Visual TouchPad is single user. Wilson [16] uses 3D sensing on tabletop systems, but for the specialized purpose of building terrain models. Similarly, Portico [1] and Bonfire [10] both provide nomadic tangible operation but Bonfire is essentially single user and Portico restricts space using a tablet for display. FACT [12] employs IR pen tracking to merge paper with a digital desktop but is aimed specifically at text markup rather than a general tangible system. Kolab shares with Carvey's [4] system, user-appropriated but unmodified objects as tangible tokens, but does not rely on weight. However, our system is designed for a collaborative context and has a greater expressivity, since it can use the 2D position and orientation of the tangibles as input rather than a token's presence or absence.

The Kolab technology and interface

The 'Kolab' prototype is a computer vision based system using a single notebook computer and an RGB-D camera (Microsoft's "Kinect"[17]) mounted on a tripod to create a multi-user touch-table upon which everyday items can be introduced, imparted with value, and arranged (see fig 1). The Kolab system has been developed to track 3D objects across a work-surface and also to provide hand tracking and a portable touch table in a form available on any surface. Kolab can also identify a range of hand gestures, giving it considerable flexibility.

In order to support augmentation-free artifacts, Kolab's object recognition and tracking are based on the 3D contour and centre of mass of each artifact. Initially the system builds a pixel-by-pixel model of the empty table using the RGB and depth values; this is then used to provide a map to segment out 'new' items introduced onto or above the surface. To help remove noise, items below an area threshold are removed. A list of tracked items is maintained. Using a 'one-nearest neighbor' search of the location of the centre-of-mass, items are tracked across subsequent frames. The depth buffer is used to identify user arm/hands occluding table artifacts. Logic is applied to account for occlusion, taking care not to remove items from the tracking list if they happen to be occluded.

The Kolab system uses an overhead projector co-mounted with the camera and 3D depth sensor. Normally this is used to project object annotations and other information onto an appropriated surface. However, for the purposes of our experiment, a separate monitor to one side of the table was used to replace the projection, with the camera view mixed in. This step was taken to facilitate the identification of a particular type of event: namely when users had to consult an annotation to check the meaning of an artifact (see figure 1), and provide a base line for use with future projector based labeling. This created the worst conceivable TUI case

The experiment, task and user interface

The objectives of the experiment were five-fold. Firstly, to test if users were able to complete tasks using tangibles appropriated from the environment. Secondly to test the impact of low 'representational significance' for nomadic tangible systems (as noted above). Thirdly, to assess users' willingness and ability to improvise and appropriate artifacts from the surrounding environment. Fourthly to investigate what kinds of artifacts users selected, and the possible implications these choices might have on the design of more sophisticated recognition systems. The final objective was to investigate which kinds of artifact,

Figure 1. Pilot Kolab desktop set up

if any introduced semiotic difficulties in performing the task (i.e. uncertainty as to what was represented). We have itemized these objectives concisely for brevity. Below, we elaborate and reflect on how they inter-relate.

The task was a standard collaborative training exercise called 'Lost at Sea'. The game required pairs of participants to agree on what items they would hypothetically take into a lifeboat to survive a sinking ship, subject to a weight limit. The projected display (seen, for this experiment, on a separate monitor) uses lines to divide the surface into a control area and an application area. The application area is itself divided into two sub-areas for the game: a 'take' area and a 'leave' area, indicating items to be taken into the lifeboat or left behind respectively. In order to import some fresh tangible to represent a new item to be taken or left, the tangible is initially placed anywhere in the application area. The control area is then used to bind this tangible to an item, as follows. The control area contains a scrollable list of all items. A dedicated control artifact placed in the control area can be slid to indicate any item from the list. Any such item may be selected using a 'double tap' hand gesture. The newly bound tangible is then duly annotated with both the name of the item it represents and its weight. The sum of the weight of kept items is also displayed, this was vital as it was one of the game aspects, along with final score, unavailable with out digital intervention .

Ecological validity was important so the experiment was conducted in a small multi-purpose meeting/training room of a medium sized Non-governmental organization (NGO). One of the office tables was used as the work surface providing a multi-user work area of 80cmx80cm, with the

MS-Kinect at a height of 69cm above the table. In order to provide a suitable contrast between the desk surface and tangible artifact the table was covered with a cloth. In order to control the lighting a large nearby windows were covered by drawing the curtains.

The work area was seeded with some items found in the room, mainly tea making items (tea bags, sugar bags, napkins, coffee cup tops & holders) as well as paper from a printer and notebook pages, a sachet of energy gel, a toy building block, train tickets and corporate postcards. Users were also encouraged to introduce their own artifacts.

Participants were volunteers from several departments of the NGO (a mix of technical and non-technical staff, all computer users) and the study was conducted at the participants' workplace over 3 days with 24 users (6 female, 18 male, all of working age) in 12 sessions. Each session lasted about 50 minutes and involved 2 participants; each participant was involved in one session only.

Each session was structured identically and video-recorded; After a brief introduction, participants were given a 10 minute tutorial to familiarize themselves and ask questions of the researcher; they then played the game twice, each time for about 10-15 minutes with the system in one of two configurations – one using a tangible 'slider', the other using free-air, above table gestures. A series of open questions were asked at session end, and a usability survey completed offline. Twenty-one user surveys were returned. The survey included ten questions from SUS, including two questions asking participants to rate on a Likert scale their assessment of the learnability of the system.

RESULTS

Given the low representational significance of the improvised tangibles, and the previously supposed importance of this factor [4, 13], it is perhaps surprising that all of the participant pairs were able to complete the task. This suggested that the weak representational significance and the lack of projected augmentation did not hinder use of the system. Thirteen (54%) of our participants, and nine (75%) of the sessions, introduced new (non seeded) tangibles into the system during the experiment, suggesting that appropriation of artifacts is natural to users. Due to limitations of the pilot image processing system, some items such as mobile phones were not reliably recognized and were not then used beyond the familiarization session. Ignoring the limitations of the pilot image processing system, Table 1 itemizes the items appropriated in the test or the familiarization session. From the third column we see that the most popular items were unsurprisingly those seeded to the participants. Yet not all the seeded objects were used equally. Coffee cup tops were a favorite (used in 100% of sessions), this may have been due to the size, tangibility and perceived disposability. Some artifact types (e.g. paper scraps) were sometimes used more than once to represent different tokens. The fourth column reports the percentage of all multiple usages. Paper scraps were used more frequently than any other artifact type, possibly because they could be more easily differentiated. This was emphasised by an (unsuccessful) attempt to stack coffee cup tops. Pens and pencils were the most introduced new artifacts echoing Bereton and McGarry's [3] finding where pens were used for mechanical analogies). Introduced items show a great deal of heterogeneity suggesting that future recognition systems should be very general.

Artifacts	New	Session Used %	Total use %	Vision issues
Coffee Cup tops		100	17.42%	
Tickets		92	7.67%	
Paper scraps		83	34.15%	
Sachet gel drink		83	4.53%	
Postcard		7	5.92%	
Cup Holder		67	5.57%	
Tea Bags		58	5.92%	
Sugar Sachets		58	6.27%	
Pencils and Pens	YES	33.33%	2.44%	Y
Napkins		25.00%	1.39%	
Mobile phones	YES	16.67%	0.70%	Y
Folded paper	YES	16.67%	1.74%	
Toy Brick		8%	1.05%	
Jewelry	YES	8%	0.35%	Y
Teaspoon	YES	8%	0.35%	Y
Paper animal	YES	8%	0.70%	
Origami cube	YES	8%	0.70%	
Coffee Cup	YES	8%	1.05%	Y
Cigarettes	YES	8%	1.05%	Y
Toy Camera	YES	8%	0.70%	
Coffee cup tops	YES	8%	0.35%	
Bottles/Drink containers	YES	0.00%	0.00%	
Notebook	YES	0.00%	0.00%	
Spectacles	YES	0.00%	0.00%	

Table 1. Artifacts introduced as tokens

It should be noted that the artifacts appropriated in this context have some significant omissions. For example, no user introduced money, wallets, purses, credit cards, security passes or business cards as tokens. We conducted a survey of the participants and found that 90% of the participants who responded generally carried a security pass, 63% money, 54% wallets/purses, 45% credit cards and 18% business cards. Carvey et al. [4] reported solo users introducing these items, suggesting that artifacts chosen in collaborative contexts may differ from solo users in a private context. In our study, only in one case did a user introduce a personal item of jewelry (a cufflink) and the least-used seeded item was the toy brick (an ideal tangible) suggesting that appropriated items in this context tend to be disposable or with no clear personal significance.

Our prototype recognizer did loose objects as shadows cast by users' hands changed object lighting without altering depth so disturbing the depth filter. Despite this in 100% of sessions, participants took advantage of the parallel nature of the tabletop and interacted with the tangible artifacts at the same time; common patterns of usage included the participants agreeing on a course of actions (e.g. "You get the rum and I'll get the fishing kit") subsequently carried out in parallel or during discussions, this suggests that the system acted to facilitate the collaborative process. Users regularly referred to the monitor (presumable to check weight limit) as users discussed the solution. Despite loosing objects users generally responded positively to the tangible interaction process. As one user said '*as a kinaesthetic learner [the idea of moving things] worked well for me*' [session #7]. In response to the question '*I thought the system was easy to use'* the prototype received scores of 3.3 (Likert scale 1...5). In response to the question '*I would imagine that most people would learn to use this system very quickly*' the system got scores of 4.05.

CONCLUSION

For a mobile or nomadic general tangible based system to be useable and supportive of impromptu collaboration we looked to design a system which would permit a very wide range of objects to be appropriated and used as tangibles. The technical approach of Kolab, using a fusion of depth and image sensing has shown it is possible to use arbitrary marker less, tag-less objects as tangible tokens. Future work will improve on the recognition of highly reflective objects and others mentioned in the vision issue column.

While for a single use tangible system it remains unarguable that the design of tokens for the right kinds of representational significance can simplify the comprehension of a tangible user interface (TUI), we have presented evidence consistent with the hypothesis that a system is not unusable when users improvise the tokens them selves for short periods. Given a number of sources in the community we were initially highly doubtful if such a system would be usable in multi user collaboration. Our work shows that a tangible system is well regarded by users as useful even in the worst case without projected on object labeling. We believe it is the tokens physicality and the process of improvisation (selection, introduction, use) that partly over comes of the aspects of the representational deficiencies. Finally despite misgivings that may emerge from the literature we have opened the way for the design of nomadic improvised tangible based TUIs.

REFERENCES

1. Avrahami, D., Wobbrock, J.O., and Izadi, S. Portico: tangible interaction on and around a tablet. *Proceedings of the 24th annual ACM symposium on User interface software and technology*, (2011), 347–356.
2. Brereton, M. and McGarry, B. An observational study of how objects support engineering design thinking and communication: implications for the design of tangible media. *Proceedings of the SIGCHI conference on Human factors in computing systems*, (2000), 217–224.
3. Brereton, M. and McGarry, B. An observational study of how objects support engineering design thinking and communication: implications for the design of tangible media. *Proceedings of the SIGCHI conference on Human factors in computing systems*, (2000), 217–224.
4. Carvey, A., Gouldstone, J., Vedurumudi, P., Whiton, A., and Ishii, H. Rubber shark as user interface. *CHI'06 extended abstracts on Human factors in computing systems*, (2006), 634–639.
5. Fitzmaurice, G.W. Graspable user interfaces. 1996.
6. Holmquist, L., Redström, J., and Ljungstrand, P. Token-based access to digital information. *Handheld and Ubiquitous Computing*, (1999), 234–245.
7. Hornecker, E. and Buur, J. Getting a grip on tangible interaction: a framework on physical space and social interaction. *Proc. of the SIGCHI conference on Human Factors in computing systems*, (2006), 437–446.
8. Ishii, H. Tangible bits: beyond pixels. *Proceedings of the 2nd international conference on Tangible and embedded interaction*, (2008), xv–xxv.
9. Kaltenbrunner, M. and Bencina, R. reacTIVision: a computer-vision framework for table-based tangible interaction. *Proc of the 1st conference on Tangible and embedded interaction*, (2007), 69–74.
10. Kane, S.K., Avrahami, D., Wobbrock, J.O., et al. Bonfire: a nomadic system for hybrid laptop-tabletop interaction. *Proc. of the 22nd ACM symposium on User interface software and technology*, (2009), 129–138.
11. Kaplan, F., Do-Lenh, S., and Dillenbourg, P. Docklamp: a portable projector-camera system. (2007).
12. Liao, C., Tang, H., Liu, Q., Chiu, P., and Chen, F. FACT: fine-grained cross-media interaction with documents via a portable hybrid paper-laptop interface. *Proceedings of the international conference on Multimedia*, (2010), 361–370.
13. Malik, S. and Laszlo, J. Visual touchpad: a two-handed gestural input device. *Proceedings of the 6th international conference on Multimodal interfaces, October*, (2004), 13–15.
14. Shaer, O. and Hornecker, E. Tangible User Interfaces: Past, Present, and Future Directions. *Human–Computer Interaction 3*, 1-2 (2009), 1–137.
15. Wilson, A.D. PlayAnywhere: a compact interactive tabletop projection-vision system. *Proc. of the 18th ACM symposium on User interface software and technology*, (2005), 83–92.
16. Wilson, A.D. Depth-Sensing Video Cameras for 3D Tangible Tabletop Interaction. IEEE Computer Society (2007).
17. Microsoft Kinect SDK. http://bit.ly/yXKLW5.