Tangible Interfaces for Download: Initial Observations from Users' Everyday Environments

Abstract
Tangible user interfaces (TUIs) have been promoted and discussed in the HCI community for 15 years. Most reported TUIs are research prototypes, available in laboratories or museums. This paper reports an attempt to understand the impact of TUIs in users' everyday environments through a low-cost, simple set-up tangible interface for music that can be freely downloaded from a website. The system requires only a regular computer, a webcam and a printer – the physical parts of the interface can be folded out of ordinary paper. Logging interaction with the interfaces and analyzing content posted by users on the web we observed that the TUIs were accepted as normal: just interfaces to make music rather than esoteric systems.

Keywords
Tangible user interface, large-scale user observation, user generated content, d-touch, music sequencer.

ACM Classification Keywords
H5.2. Information interfaces and presentation (e.g., HCI): Evaluation/methodology.

General Terms
Human Factors, Experimentation
Introduction

Tangible user interfaces (TUIs) have been promoted and discussed in the HCI community for 15 years. In TUIs physical objects are used for the control and representation of digital information [6,8], similarly to how icons are used in graphical user interfaces for the same purpose. Proposed areas of application span from education [5,15], to creative expression [9,11].

Most reported TUI systems have the nature of research prototypes, available in laboratories, or interactive installations on display in museums. This is partially due to expensive or complex set-ups, involving, for example, custom-built electronic sensing systems [11] or retro-projected surfaces [9]. As a consequence, most accounts of TUI usage to date have been restricted to controlled settings [15] or short-term interaction [5]. The suitability of TUIs for application in users’ everyday environments, such as homes, offices and schools, as well as people’s reaction beyond the initial stage, remain open questions. At the same time, a recent trend in HCI is raising the attention towards Web 2.0, user-generated content (UGC) and online communities as resources to gather research data [2,10,14].

We argue for this approach to be extended further, making tangible and embodied interfaces available through the Internet to remotely study their adoption. As a first example, in this paper we report an attempt to understand the impact of TUIs in users’ everyday environments through an Internet-based observation of volunteer users. We developed Audio d-touch, a collection of low-cost tangible interfaces for musical composition and performance, and made it freely and fully available for download from our website. Audio d-touch requires only a regular computer, a webcam and a printer. The physical interactive objects can be folded out of ordinary paper, with visual markers printed on it, or made by gluing printed labels onto existing objects. The webcam, hanging from a desk lamp or other improvised stand, tracks the interface items using the d-touch marker recognition system [3]: the objects position is mapped to music synthesis parameters. We logged interaction with the interfaces and analyzed content posted by users on our own and other websites to observe and evaluate how people relate to such novel systems. Beside predictable enthusiasm for the novelty factor of the interfaces, most users’ comments focused on specific aspects related to the musical nature of the applications and their integration within existing ecologies of software music tools, indicating that the TUIs were accepted as normal: just interfaces to make music rather than esoteric systems.

The next section provides an overview of related work, followed by a description of the Audio d-touch applications and their simple installation. The on-line promotion, the remote observation of the system and the analysis of user generated content are then reported, followed by a discussion and the conclusion.

Related Work

In the past fifteen years a number of researchers have worked on tangible user interfaces. Fitzmaurice et al. explored the possibilities of “graspable user interfaces” in 1995 [6]; Ishii and Ullmer introduced the term of tangible user interface in 1997 [8]. Audio d-touch follows the paradigm of tangible user interfaces, in particular in the field devoted to musical applications. Musical TUIs are already established in research and commercially, as proven by Reactable and Audiopad.
These instruments generate sounds through the interaction of the different blocks and their motions. A large number of other amateur and professional built musical TUIs fit into this category with one of the most popular being BeatBearing [1]. Other musical systems include touch based instruments on table-tops such as the Microsoft Surface or Jeff Han’s prototype [7]. The main difference with Audio d-touch lies in its low-cost and easy do it yourself properties. Instead of just watching prototypes in labs, museums or expensive recording studios, users can try it at their home. The other main difference with these systems is that the Audio d-touch applications are developed as a tool to remotely observe and track people usage of TUIs in their own settings, enabling the first large scale test of musical TUIs outside a laboratory.

Thanks to the availability of ARToolkit and other toolkits as libraries for Adobe Flash, recently a number of augmented reality demos started to be available on the Web. Users can print markers, point a webcam to them and see them on their monitor augmented with 3D objects. However, many of these projects seem purely targeted at creating a wow-effect for advertising of products and services, rather than exploring new interfaces for actual applications. Moreover, no report was found of studies assessing the users interaction with these systems. In contrast our aim is to use the Web to distribute complete applications that can be controlled using a tangible user interface, and study how real users interact with them.

Audio d-touch Drum Machine and Sequencer

Audio d-touch is a small collection of applications for real-time musical composition and performance. It includes a drum machine and a sampling sequencer. Both are controlled by physically arranging a set of grasprable interactive blocks on a flat surface (e.g. a table-top). We refer to the surface where the interaction takes place as the interactive surface or the board. The blocks position relative to the surface, tracked through a webcam, controls the parameters of the audio synthesis applications. Both the blocks and the interactive surface are marked with ordinarily printed pieces of paper, containing d-touch markers [3] used to convey interaction cues to the users and to make the automatic tracking easier. Four markers are placed at the corners of the interactive surface, for calibration, so that the interface behaviour is tolerant to small camera movements. The interactive objects can be built as small foldable boxes made of cardboard, or simply by attaching the labels to any small objects available, such as nuts, small chocolate bars, candies, toy bricks... To position the camera so that it correctly observes the interactive surface it can be attached to a desk lamp, or standard tripod. Low cost and easy set-up were factors driving the design of the project: only off-the-shelf consumer-grade hardware and printed paper.

For ease of construction the interactive surface is normally delimited by an A4 piece of paper, which defines a block size of approximately 2.5 by 3 cm, as shown in Fig. 1. However, the system can be scaled to practically any size, as long as interactive surface and blocks are scaled by the same factor and everything is in the field of view of the camera.

---

[2] For example: www.megabaile.com or www.bmw.co.uk/z43d or www.mini.de/webcam
Sounds and other objects are represented by the graspable blocks, while actions (e.g. play, record, ...) are represented by flat areas of the interface, we call these active areas. The relative position of a block inside an active area can determine variations of how the action is applied to the object. This description is based on Schneiderman’s action-object paradigm [13]. We think that in the context of tangible user interfaces associating objects to graspable items and actions to flat elements can reduce ambiguity. In other words, the Audio d-touch interfaces are defined through spatial mapping: they are operated by placing the blocks in specific positions. From the audio point of view, both applications are loop-based sequencers. Sounds get reproduced when hit by an invisible virtual cursor that periodically scans the interactive surface from left to right.

In the d-touch drum machine object-blocks represent drum hits and the entire interactive surface is covered by a play active area: when blocks are placed on it the corresponding drum hits are played. The x-position corresponds to the time instant within the loop when the sound is played, while the y-position determines the type of drum, 11 in total, as shown in Fig. 2. Two types of blocks are available: normal hit or louder hit.

For the d-touch sequencer the interactive surface contains 3 active areas: a record area, a store pattern area and a larger play active area divided into two identical tracks, as illustrated in Fig. 3. Eighteen block types act as sound container objects, in the play area the x-position determines the playback instant within the loop (as in the drum machine), while the y-position relative to the track determines the playback volume. The rotation of the block is mapped to the playback speed, colour triangles on the blocks indicate the forward playback direction. When a block is placed in the record area the live input to the soundcard gets recorded "on it". Similarly, when a block is placed in the store area the content of the playback areas gets assigned to it. The size of the recording and storing areas are such that they can only contain one block at the time.

TUIs to the Masses
The Audio d-touch applications have informally been tested by a number of musicians and non-musicians across a wide variety of situations [4]. We generally received enthusiastic reactions, people were intrigued by the system, for its unusual aspects but also for its playfulness, simplicity and low cost nature. Musicians particularly appreciated using an electronic instrument without a video monitor and compared it to audio effects pedals. This informal positive feedback, demonstrating that the instruments were potentially appealing to a wide audience, encouraged us to try and arrange a large scale observation. The low cost nature of the system was key in deciding to make it available for download from a website. The decision was also influenced by the general popularity of Reactable [9].

To observe how users interact with Audio d-touch, and in turn understand how they relate to its type of interfaces we used remote logging. We decided to collect the minimum amount of information that would let us reconstruct the interaction while limiting the invasion of the privacy of our users. The coordinates of all markers detected by the system, frame by frame, and (only for the sequencer) the audio clips recorded

---

3 Videos of the concerts on http://d-touch.org/audio/concerts
by the users were transmitted over Internet to a server, marked with a timestamp and stored there in a private database. To be able to identify repeated usage patterns we requested users to register on the website and employ the same username and password when the application is launched. During the registration process users were asked about their age, gender, occupation, whether they play any musical instrument and how frequently and whether they had any knowledge or practice about tangible user interfaces, and if so how. An overview of the information collected is in the side box "Collected Data". A number of measures were taken to respect the privacy of our voluntary users. The remote logging operated by the system is clearly stated in the conditions of use that users have to accept before downloading the software. Users can have their data removed from our storage at any point.

To distribute the system we created the d-touch.org website and we designed a kit in the form of cut-along-the-lines graphics distributed in PDF format. The purpose of this kit was mainly to help the project promotional communication – we wanted to convey that Audio d-touch can be simply and precisely reproduced following straightforward instructions and easy-to-find materials. We released the applications, the graphic files for the physical parts of the interfaces, video and photo documentation showing the instruments in action and how to set them up. Initially only the drum machine was released, while the logging system for the sequencer was still being completed. The videos were also posted on YouTube, to facilitate their diffusion.

**Analysis**

**Reconstructed Videos**

As it was not possible to gather meaningful numerical data out of the reconstructed videos, we analyzed the content in a qualitative way. As we have only analyzed a subset of the videos, due to the overall length greater than 90 hours, we provide just trends, not percentages.

A majority of the users managed to explore the interface and to eventually produce a basic rhythm. By exploring we mean moving around blocks on the calibrated active area, and we define a basic rhythm as a meaningful pattern of more than four blocks. For this exploration to make sense the system needs to be calibrated, i.e. the four markers on the corners of the active area need to be detected most of the time, which means that the camera has to correctly point to the sheet. Relatively few users tried to go beyond the basic rhythm step and produce advanced rhythms with many blocks: only 20% of the sessions contained more than 10 blocks.

Another phenomenon observed in the reconstructed videos is that, in some sessions, markers are recognized intermittently, which negatively affects the user experience, as the layout of physical blocks may not correspond to the audio generated by the applications. Given that the algorithm has been massively tested through more than 5 years in a variety of conditions, we suppose that this issue is related to severely poor lighting conditions. After this observation we realized that we did not provide users specific guidelines about lighting, and that the feedback that the applications provide about proper lighting is difficult to interpret.
Generally users were able to create rhythmic patterns, except when the recognition was intermittent, or the camera unstable. Under these conditions, users seemed to engage in a basic exploration of the interface: they would place one or few blocks on different position, as to explore what happens.

**User Generated Content: Videos and Photos**

We analyzed the user generated videos and photos to find out how and where real users setup the Audio d-touch instruments. First of all we observed that some people preferred to use markers on flat pieces of paper, without folding it to make up 3D graspable blocks. It was interesting to see that sometimes entire uncut strips of markers were used as blocks to construct repetitive patterns. On the opposite end, we observed one user who built blocks out of wood to create a very polished setup. The cardboard stand that we offered for download was used to set up Audio d-touch on tripods or hung from shelves. In some cases, the paper board was raised to bring it closer to the camera.

In all the available material Audio d-touch was set up on desks, and never on other casual surfaces, such as coffee tables or on the floor. The desks used were often rather small and cluttered by different items, so the interactive board was rarely placed straight in front of the user (following the direction defined by the type), instead it seemed positioned just as space allowed. The interfaces were generally setup in small rooms, bedrooms or individual offices. The hardware used for the system was varied: laptops and desktop computers, Windows and OS X. The audio output ranged from laptop built-in speakers to professional hi-fi systems.

In 2 instances we observed setups that we found very interesting in that they appropriate the technology in radical ways. In one setting, a user printed a board around 8 times bigger than the standard A4, and she affixed it vertically to a magnetic whiteboard. The blocks were made by gluing the printed d-touch labels to office magnets. In the video the user explains that such big and vertical setup more easily allows collaborative usage; two fellow students are filmed while they create a drum sequence. In the second case we witnessed a setup of the drum machine without paper or other physical blocks. The webcam is pointed to the computer screen, where a graphic design software (perhaps Adobe Illustrator) is used to display the interactive board and icons representing the markers, all operated through the standard computer mouse. This subverts completely the original aim of the system!
To analyse the text generated by our users we adopted an approach inspired by grounded theory [2,13]. The analysis started by categorizing the material at the sentence level through open codes. Initially 50 open codes were used, later grouped in 6 broader categories: physicality of the interface, audio, technical, improve and extend, field trial related, generic.

**INTERFACE PHYSICALITY AND EASE OF USE**
From sentences related to the tangible interface we gathered that users appreciated the physical representation of sounds: "When I showed my good bands and songwriters the setup, they loved the realness and ability to touch and move something real to make the sound" or "Software doesn't have to mean virtualizing everything and letting go of physical objects. On the contrary, it can create all sorts of imaginative, new ways of mapping musical ideas to the physical world. And that's how we wind up with a walnut drum sequencer."

Others appreciated the ease of use of the interface: "..the interfaces are extremely user friendly and the simple design of the projects: It's easy and it's really fun. All other examples of similar technologies involved being a computer genius and the step by step instructions are just pain awesome for everyone, this is the right way to do this kind of things, by making them available and enjoyable to everyone."

**AUDIO**
A good number of entries were related to the audio nature of the system. As exemplified by the following quote posted on the forum, audio proved to be a playful and engaging domain of application: "Everyone in this house has now put together a radio-worthy beat by pushing little scraps of paper around under a webcam."

Other users suggested that Audio d-touch could be a good tool to teach music in school, or that they built the system for their children to enjoy the easy music generation with them.

Audio was also the main topic of criticism directed to the system. The limits and incompleteness of the music synthesis parts of the applications are probably the main reason why users stopped using Audio d-touch early. Several users requested the drum machine to send MIDI signals or to be able to load custom samples. Especially for musicians these two restrictions limited the usage.

**TECHNICAL**
Under the technical category we gathered bug reports, problems with the software setup and comments about technical aspects. The entries in this category show a great interest by our users in trying and understanding the software. Users who posted this kind of content generally showed technical competence.

**IMPROVE AND EXTEND**
Some users particularly enjoyed the actual building phase: "I spent a happy half an hour cutting out the shapes and putting the little boxes together."

Nobody reported the setup as difficult or bothersome, instead some users proposed much more complex structures: "Currently I made a stand with 2 pvc pipes, 1 pvc L joint, and a metal flange for the base, and an adapter to go from the pvc pipe to the metal flange. I then threaded the camera cord through that, cut notches out at the top to secure the webcam ‘clip’, and the height is just enough to accommodate the board being on a 8.5 by 11 inch paper board."

<table>
<thead>
<tr>
<th>Table 1: Audio d-touch usage data, gathered from the interaction logs. In brackets the standard deviation of the average values.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Avg session length, minutes</strong></td>
</tr>
<tr>
<td>Sequencer 8.75 (10.12)</td>
</tr>
<tr>
<td>Drum 6.75 (8.05)</td>
</tr>
<tr>
<td>Overall 7.34 (8.76)</td>
</tr>
<tr>
<td><strong>Avg no. sessions per user</strong></td>
</tr>
<tr>
<td>Sequencer 3.33 (3.13)</td>
</tr>
<tr>
<td>Drum 5.45 (6.35)</td>
</tr>
<tr>
<td>Overall 5.55 (6.34)</td>
</tr>
<tr>
<td><strong>Max no. sessions per user</strong></td>
</tr>
<tr>
<td>Sequencer 15</td>
</tr>
<tr>
<td>Drum 30</td>
</tr>
<tr>
<td>Overall 39</td>
</tr>
<tr>
<td><strong>Avg blocks in session</strong></td>
</tr>
<tr>
<td>Sequencer 1.80 (1.71)</td>
</tr>
<tr>
<td>Drum 3.06 (3.75)</td>
</tr>
<tr>
<td>Overall 2.69 (3.33)</td>
</tr>
<tr>
<td><strong>Max no. sessions per user</strong></td>
</tr>
<tr>
<td>Sequencer 15</td>
</tr>
<tr>
<td>Drum 30</td>
</tr>
<tr>
<td>Overall 39</td>
</tr>
</tbody>
</table>
FIELD TRIAL RELATED
Some people complained about the remote data logging and the registration requiring from them too much data. However, we generally received appreciation for this research. Often the fact that we were observing the usage was accepted because the application was free to download. Some users enquired about the possibility of buying a version of the applications that does not require Internet connection, because this made it difficult to use the system on stage or in studios.

GENERIC COMMENTS
A number of comments do not appear related to actual usage of the system. This content is more generically about emotional and initial reactions to seeing the project on the web. Even though this material is less informative than actual usage feedback, we report it because it shows a large general interest for the Audio d-touch system and, more in general, for tangible user interfaces that can be freely downloaded. Numerous comments, especially on blogs and on Twitter, expressed surprise and enthusiasm for the project, a wow-effect: “Sci-Fi??? No more! You can actually build it!” or “This is an awesome UI.” Many Twitter users posted just a link to our website or to other websites featuring our project. A very small part of the Twitter content was about personal feelings or experience about the project, but it was always superficial.

Discussion
The analysis of the reconstructed interaction videos shows that the simple set-up works. Many users managed to get the system to work perfectly with the minimal amounts of instructions provided from our website, demonstrating that the low cost nature of the system was effective for spreading an experimental technology to a large number of users in their own environments. The general concept of interacting by laying out blocks over the board with the camera pointed to it is clear. However, from the reconstructed video analysis, we could see that in some sessions markers are recognized intermittently, even to the extent that calibration could not be performed. This was most likely due to poor lighting conditions and to inadequate camera position and orientation, given also that very little information was provided to users about optimal lighting conditions. Moreover audio d-touch currently provides little feedback about correct setup.

Most users explored the interface and produced basic rhythm, but very few used it regularly with more advanced patterns. This is probably due to the fact that the instrument is seen mostly as a toy, as pointed out in some of the comments: many users did not considering audio d-touch as a real music-making software, rather as a way to have fun making simple music beats or to explore a new type of interface. The trends found in the interaction logs (Table 1) supports that users generally perceived audio d-touch as toy. The more playful drum machine has been used more than the more complex sequencer (65% of usage) and the time spent on Audio d-touch is compressed in less than 2 days for the 79% of users. As a trend, fewer blocks were used in the sequencer and its sessions were longer, suggesting that this application allows the creation of more interesting pattern with fewer blocks (thanks to the record and store functions).

From the observations of the video sessions it emerged that almost all users explored the functionality of the new system to different degrees. In some cases the exploration was limited to the assessment of the basic
functionality of the system: placing one or few blocks on the board (with the system being calibrated or not) users experienced the effect of having them in different positions, without attempting to create rhythmical structures. In other cases, the exploration was more directly targeted at exploring the ability of the system to create complex rhythmic patterns. The high number of people registering on the website compared to the number of people successfully trying out the interface may be due to several reasons, and it is difficult to make precise guesses; we suspect that incompatibility with specific hardware platforms may be one of them.

Several users made good points about missing features or technical problems related to the application, but no one had specific negative comments about the tangible interface. We have in general very few comments about it, even if it’s a completely new interface. Only a small minority of our users reported previous experience with tangible interfaces, nevertheless we observed, from quantitative and qualitative data, a rapid learning especially for those with musical background, showing that the interface is almost natural to users.

The user generated content showed several examples of user appropriation. Two teachers who wanted to use it in their classes, one to teach music to young children and the other to teach game design, they also changed the Audio d-touch setup to better fit their purpose, even though this involved more complex construction. Others, who wanted to use the d-touch applications for their music band, told us that they wanted to make the board in wood and use heavy blocks. Some users contacted us about building a large scale version of the system to be used in festivals or performances. In one case d-touch was used even without a printer on a computer screen, making a virtual tangible interface.

In summary, we argue that the web distribution and user adoption of the audio d-touch tangible user interfaces was a success, even though the applications were perceived more as toys than as proper musical instruments. The quantitative data from the logs show that, beside technical problems, a large number of users were successful in interacting with the tangible interfaces and exploring their functionality. Very few comments were made about the user interface itself, and all of them were positive. As discussed above, we interpret this lack of comments as evidence that users found the interface “natural” or “obvious”, despite the fact that very few of them reported having experienced a tangible user interface ever before. Probably this circumstance is partially due to a considerable presence of tangible and multi-touch interfaces in popular media in recent years. Finally, the multiple cases of user appropriation indicate a strong interest and advanced understanding of this technology.

**Future Work**

Based on the analysis of the reconstructed videos, our top priority for future development is the improvement of feedback about the system correct setup. We are also looking into ways to allow people to easily share their compositions, to foster the development of a community. Compositions downloaded from d-touch.org could be printed and used as a physical basis for further interaction. Redeveloping the system using an interpreted language such as Adobe Flash could reduce hardware compatibility issues. If the performance penalty is not a show stopper, such an
approach would make Audio d-touch available to an even larger audience directly through the browser.

In the future we are interested in developing new low-cost tangible interfaces to include also visual output. This could be implemented through low-cost pico-projectors (to be hung next to the webcam) or using LCD screen placed horizontally, in place of the printed interaction board. While using LCD monitors is exciting because of their ubiquitous availability it also raises some technical problems (due to constant backlight).

Conclusion
This paper reported an observation of tangible interfaces in users’ everyday environments. We observed users through interaction logs and analyzed the comments they posted on forums as well as multimedia UGC they posted on websites. Despite some technical difficulties and missing features at the application level that could enhance the overall experience of the user, audio d-touch received very good responses. Few comments were expressed about the interfaces per-se and all of these were strongly positive; we noticed several examples of user appropriation. These observation show that the time is mature to distribute tangible user interfaces, also in domains different from the audio, in an inexpensive and democratic way. It’s time to bring TUIs to the masses!

References