Evaluating the Cognitive Aspects of User Interaction with 2D Visual Tagging Systems

Samuel Olugbenga King
Loughborough University, UK

Abstract—There has been significant interest in the development and deployment of visual tagging applications in recent times. But user perceptions about the purpose and function of visual tagging systems have not received much attention. This paper presents a user experience study that investigates the cognitive models that novice users have about interacting with visual tagging applications. The results of the study show that although most users are unfamiliar with visual tagging technologies, they could accurately predict the purpose and mode of retrieval of data stored in visual tags. The study concludes with suggestions on how to improve the recognition, ease of recall and design of visual tags.

Author Keywords—Mental models, ubiquitous interactivity, visual tags, social computing, HCI, evaluation, input devices, mobile interaction.

ACM Classification Keywords—H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

I. INTRODUCTION

The proliferation, portability and affordability of mobile devices with magnified computing and processing power have led to an increased use of cell phones for social, on-the-move interaction. The use of 2D visual tagging systems in particular enables individuals to interact with and collect information about objects in the physical environment. At the core of a 2D visual tagging system is (usually) a camera phone which has the appropriate decoding software and may be used to capture a visual tag. A visual tag may encode a URL (Web address) link, text, image or personal information. By connecting physical media to mobile content therefore, 2D visual tagging systems create new affordances and pathways for interacting with objects and media in the physical, educational and social environment. Although there are a number of visual tagging systems in the public domain, for example Smartpox[I] and KoolTag[II], the cognitive and psychological aspects of interacting with visual tags have not been comprehensively investigated.

II. RELATED WORK

Pertinent literature review shows that there has been increased focus on the description and deployment of tagging applications. In [6], a 2D visual code system for marker-based interaction is described. In [2], TRIP, a 2D marker technology for 3D vision-based (sensor) location- tracking system is described. The main features of TRIP – and by extension, [5], and [7] are that it is relatively inexpensive and easy to deploy. The TRIP system is composed of 2D barcodes/ringcodes and CCD cameras, and is used to determine the 3D position and identifier of tagged objects. On the other hand, Sony Japan [5] developed a 2D tagging system, CyberCode, which could specify the 3D location and ID number of tagged objects. The CyberCode system is designed for the augmented reality environment and work with stationary cameras such as those on PDAs or notebooks.

It is important to point out however that there is a dearth of research with focus on user perceptions about how to interact with 2D visual tagging applications. Literature review also does not reveal any effort to measure the psychological considerations for the deployment and use of visual tagging systems in a developing world context. However in [1], Sara Belt, et al. presents a study of an evaluation of users' perceptions about interacting with RFID and visual tagging technologies. The strengths of the [1] research were that the sample was representative of the Finnish population and the user studies was conducted in a shopping mall which is a realistic "everyday life environment". The study is also probably the first to investigate user perceptions about mobile visual tagging interactions.

Another distinction of the [1] study is the evaluation of a 2D visual tagging system that was not developed by the authors to investigate user perceptions about mobile interactions with visual tagging technologies. Research efforts in the field of visual tagging have tended to focus on the description and deployment of novel tagging applications. However, [1] involved the investigation of only one visual tagging platform - that is, Semacode. Also, the Finnish are very sophisticated in terms of mobile phone usage and adoption of ICTs. The outcome of such a study of user perceptions about interacting with visual tags may therefore not necessarily be valid in a Sub-Saharan Africa context.

III. USER EXPERIENCE STUDY DESIGN AND RESULTS

The design together with the results of the user experience study, which consisted of mental model usability testing sessions, observations and interviews, is presented in this section.

A. *Sample*

9 female and 11 male participants volunteered for this study. Most of the participants owned or used a cell phone regularly. However most of the participants had mobile phones with VGA camera capability, while only a few had camera phones of 1.0 megapixels and above (see Figure 1). In addition, none of the participants was familiar with 2D barcode systems or

[I] www.smartpox.com
[II] www.kooltag.com

had previous interaction experience with a visual tagging system. All the participants were drawn from the non-academic staff and student population of the University of Cape Town community. Furthermore, none of the participants had any physical defects that could have affected their performance during the study.

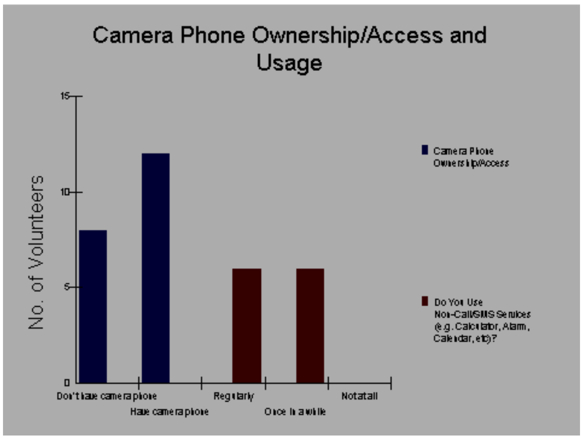

Figure 1. Camera phone ownership access and usage.

B. *Test Environment*

The Department of Manuscript and Archives (MSS)[III] at the University of Cape Town was selected as the test environment for the user experience study. The MSS was selected because it is an 'everyday life environment' where real life visual tagging application scenarios could be tested. The MSS has an extensive collection of original research material – manuscripts, photographs, et cetera - relating to the political, social, cultural and economic history of the Western Cape in South Africa. The subjects covered include art, music, education, literature and language, botany, politics and architecture. However, most of these collections are in print form. Only a few of the collections have been digitized. The collections that can be accessed online include the San (Bushman) online photographic collection which was selected for this study. Visual tags that link to the URL for the San photo album were created. The visual tags were then placed next to the print equivalent of the photographs, and participants were asked to interact with the tags.

C. *Equipment*

The equipment for this research project consisted of a Nokia 6280 camera phone that had been pre-installed with the relevant visual tag readers as well as Semacode[IV], Shotcode[V], ConnexTo[VI] BeeTagg[VII] and UpCode[VIII] visual tags – see Figure 2. These five 2D visual tagging systems were selected because of their robustness and high visibility in the field of physical world hyperlinks; and also because a Nokia 6280 handset can be used to interact with any of these systems. Due to financial constraints, only a Nokia 6280 handset was used for this study.

III http://www.lib.uct.ac.za/mss/
IV www.semacode.org
V www.shotcode.com
VI www.connexto.com
VII www.beetagg.com
VIII www.upcode.fi

Figure 2. An example of the use of an UpCode visual tag within an ad to purchase a ticket or pay bills (All a user has to do is click on the UpCode tag in the picture to initiate the transaction). Used with permission.

D. *Mental Model Testing*

When human beings are presented with a task or situation they had not encountered before, they usually fall back on a repertoire of psychologically-archived 'mental models' of how similar tasks or situations had been handled before [3] [4] (A good introduction to the concept of mental modelling in usability is available online[IX].) Symbol-reading technologies are new and so users were assumed to be unfamiliar with the 2D visual tagging paradigm. Consequently, the aim of the mental model usability testing sessions was to evaluate the perceptions of users on how to use and interact with the printed 2D visual tags with the Nokia 6280 camera phone provided. Participants were not told what to do or how to proceed; but they were given an overview of the functions and application scenarios of 2D barcodes at the beginning of a usability testing session.

The major goal of the mental model usability testing sessions was to get a verbal explanation or demonstration from users on how they were going to use or obtain information from a 2D visual tag in spite of the absence of relevant cognitive models. In addition, the determination of the mental models or perceptions that users had about how to use visual tagging systems, was dependent on the outcome of the mental model usability testing sessions. During the mental model usability testing sessions, participants were required to carry out three tasks. The tasks are:

- 2D visual tag recognition
- Inference of information stored in visual tags, and
- Mode of obtaining information stored in the visual tags.

This section details the results of the mental model testing phase.

E. *Visual Tag Recognition*

At the beginning of each of the 20 mental model usability testing sessions, all 20 participants were

IX http://www.lauradove.info/reports/mental%20models.htm and http://www.interaction-design.org/encyclopedia/mental_models.html.

shown a visual tag. The visual tag (that is, Shotcode, Semacode, ConnexTo, BeeTagg or UpCode visual tag) a participant was shown was dependent on the experimental order observed for that particular usability testing session. They were then asked whether they were familiar with, or could recognize the tag they were shown. Most participants (15 participants or 75% of respondents) were not familiar with the 2D visual tags – see Figure 3. A few participants did however spontaneously suggest that the tags looked like crossword puzzles or mathematical symbols.

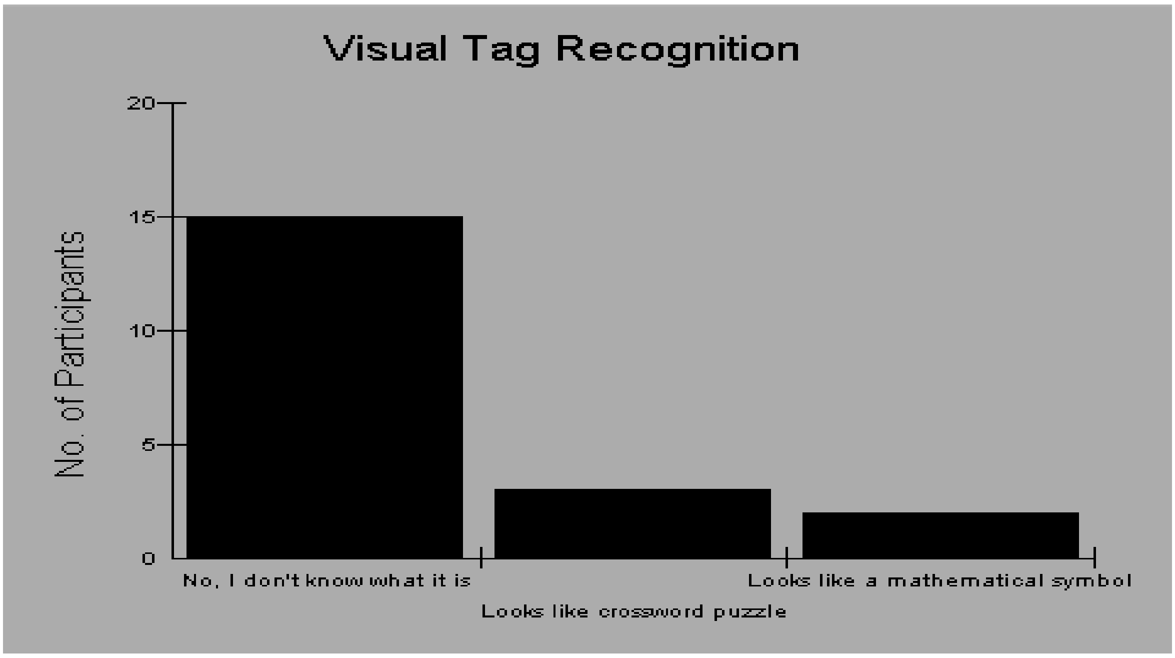

Figure 3. Visual tag recognition.

F. Inferring Data Stored in Visual Tags

The second task that participants were required to carry out was the inference of information stored in the visual tags. For this testing phase, participants were again shown a visual tag and were introduced to the concept of barcodes, using books as an example. They were then prompted for suggestions on the kind of data the visual tag they were shown could store. The visual tags for this task encoded URL links to the San photo album in the UCT MSS library[X]. Thus each of the visual tags used in the study encoded a URL link but participants were unaware of this. Moreover, the visual tags used for this testing phase were placed next to the UCT MSS library logo, and participants were told there was a link between the UCT MSS library and the visual tag. The responses of participants to this activity were varied. Most guessed accurately that the visual tags encoded resources in the UCT MSS library. A selection of the typical responses from participants about the kind of data that could be stored in tags is presented below:

- *Use for photos / store photos*
- *Where the books are*
- *The content and authors of books*
- *Books, brief content, summary of photos, authors*
- *Text, pictures, sound, programming code, cataloguing system*

G. Mode of Obtaining Information Stored in Visual Tags

Once participants had stated the kind of information they think could be stored in visual tags, they were given a Nokia 6280 phone and told that they had to use the Nokia phone to obtain the UCT MSS library data stored in the visual tag. They were then asked to show how they would retrieve the data stored in the tags, using the phone provided (Figure 4).

x http://www.lib.uct.ac.za/mss/existing/DBleekXML/Website/index.htm.

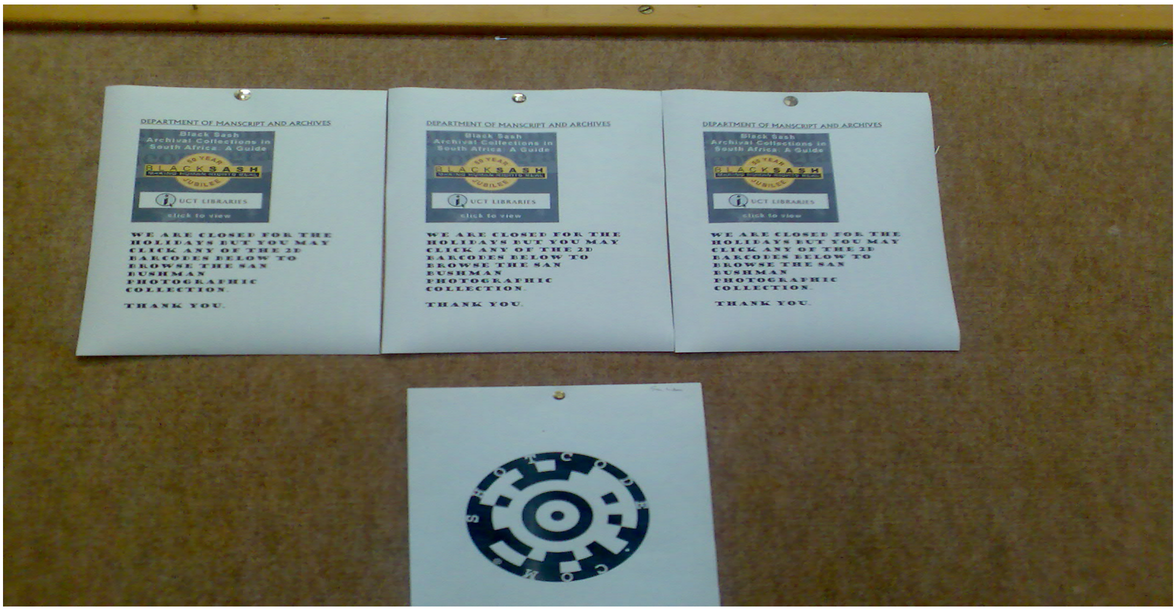

Figure 4. A ShotCode visual tag which stores the URL for the San Photo album that is on the UCT MSS Library site. Participants were asked to demonstrate how they would use a camera phone to retrieve information from the visual tags provided.

The responses to this activity were also varied. While some participants had either accurate or close guesses as to how tags may be decoded, others were completely off the mark in their responses. Some of the comments from the former group include:

If the cell phone is a recorder, take a photo – zoom in with built in camera

Take a picture of the barcode

Use special application to decipher code dots or info

A program on phone that will decode tag

Analysis of the comments from participants - as typified by the responses below - also indicate that some participants thought that visual tags could be decoded via Bluetooth, Infra red technology and/or the Internet:

Through Bluetooth, Infra red technology

Infra red technology scanning

email image to someone/org for analysis

Type in some URLs/ message on code and look it up on the Internet

Also, go to Shotcode.com that's printed on the tag

SMS the code no to a dedicated cell phone number, and they will send a Web link. Click on the phone and gets you to the Web site

If the tag has code number, google it.

H. Design Appeal of Visual Tags

As part of the interview questions at the end of a mental model usability testing session, participants were asked to rank the five visual tags (BeeTagg, Semacode, ConnexTo, UpCode and Shotcode) used in this study in terms of design and visual appeal. This exercise was carried out because visual tags serve a dual purpose: they encode information, but they also advertise a service. Consequently, the visual appeal of tags must be such that their presence in the physical environment should be able to attract both those who are familiar with their use, and more importantly, curious passersby drawn by the aesthetic or design appeal of the tags. The results show that participants consistently rated the Shotcode visual tag as having the best design in terms of visual appeal and brand recognition. It seems the circular shape of the Shotcode tag made it stand out ("Interesting picture"“; Nice image or barcode to capture"; "Easier to capture code as it is circular…"). Participants also rated both Shotcode and UpCode visual tags highly for their instant tag recognition feature - both tags have their brand names imprinted on them ("UpCode and Shotcode have text/brand names – easier to recognize and relate to"; "The Shotcode and UpCode have good graphics. They fit in the norm and would blend in with other posters"). It should be noted here that although the UpCode tags are available in colour format, only (printed) black and white UpCode tags were evaluated for visual appeal. Some participants were also impressed with the BeeTagg visual tag due to its distinctive beehive-like design. In contrast, participants were generally not impressed with the visual appeal of the ConnexTo and Semacode tags *("ConnexTo and Semacode look like technical computer code – difficult to understand").*

The ranking of the visual tags by participants based on visual design appeal (see Figure 5) shows that 47.1% (or nearly half) of participants ranked Shotcode as having the best visual tag design, while just under 3% of participants thought that Semacode visual tags have a high visual appeal.

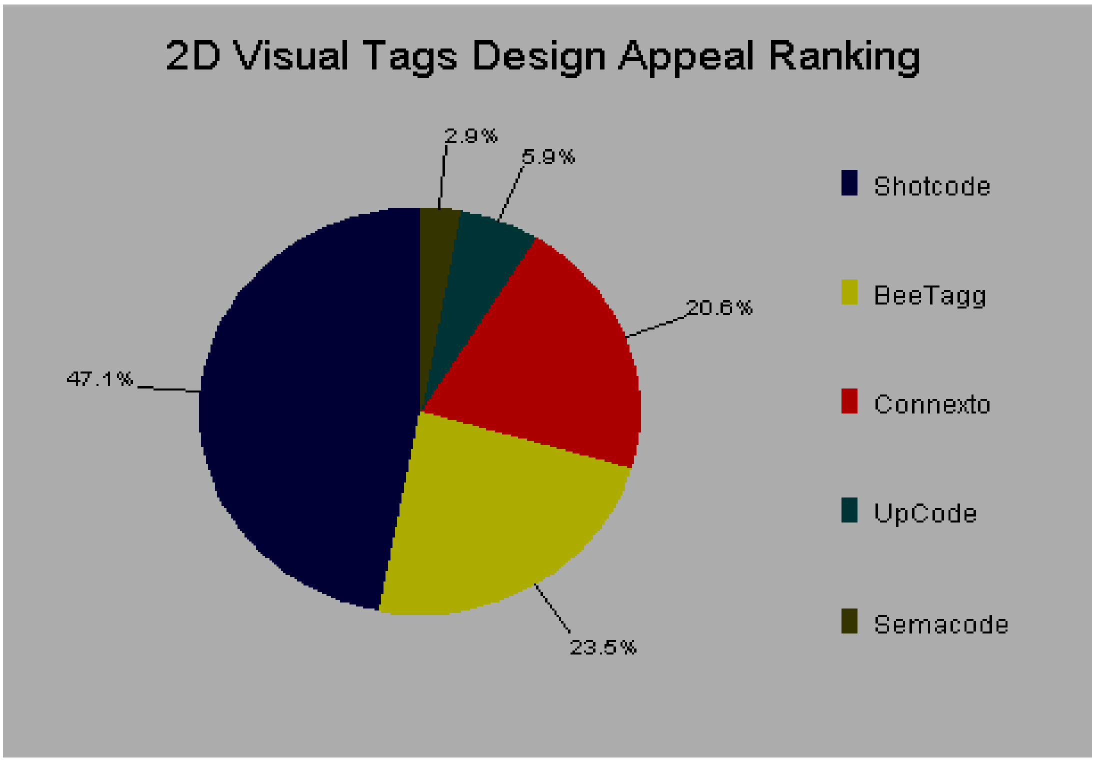

Figure 5. Visual tags’ design appeal and ranking.

I. Ease of Use of Visual Tags

Participants were also asked after the conclusion of all the mental model usability testing sessions to name the 2D visual tag system (i.e. Shotcode, BeeTagg, ConnexTo, or UpCode) they found easiest to use. The Semacode system was not included in this phase because participants did not use the system for direct 2D visual tag decoding purposes. The results of this ranking exercise (presented in Figure 6) shows that 71.8% of all participants felt that the Shotcode and BeeTagg visual tagging systems were the easiest to use.

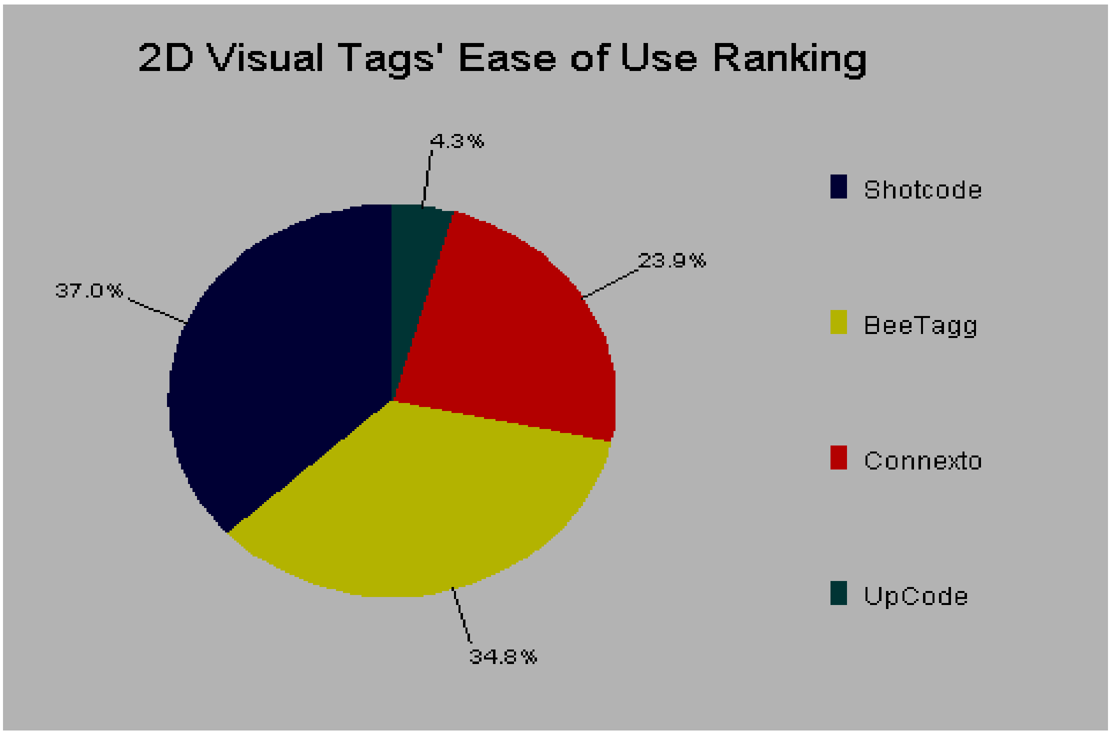

Figure 6. Visual tags’ ease of use ranking.

IV. CONCLUSIONS

This paper presents a user experience study that focuses on the investigation of user perceptions about interacting with visual tags through the use of a cell phone equipped with the appropriate software. Although none of the participants in this study had any prior experience with 2D visual tagging systems, the results of the mental model usability testing sessions reveal that most participants were able to accurately predict the type of data that would be encoded in a visual tag when shown one. Furthermore, about a third of participants were near accurate in specifying the mode of getting information from, or decoding the 2D visual tags. Therefore, from a mental modeling or perception perspective, novice users will be able to accurately infer the function of 2D visual tags when they encounter such tags in their neighbourhoods or physical environment.

Moreover, the feedback from participants shows that users are drawn to tags that are distinctive and thus have a 'come hither' appeal. But since almost all participants could either not recognize visual tags or accurately state their purpose, it becomes imperative then that attention be given to promoting the visual tagging paradigm platform to novice users or those who are unfamiliar with the concept and use of visual tags. There are three possible approaches to promoting awareness and enhancing the visual appeal and design of 2D visual tags.

First, visual tags must be designed in such a way that they must be able to 'stand out' in the physical environment. One reason why this is necessary is that the natural environment is saturated with images such as colourful posters and billboards; and so only distinctive 2D visual tags will be able to stand out from the motley of posters and adverts that are all vying for human attention. This means that visual tags should be available in multiple colours. This means also that the performance of the imaging modules or image recognition component of the current generation of 2D visual tagging systems will have to be enhanced to facilitate the decoding of multi-coloured visual tags. This is because 2D visual tag readers are currently either designed to read, or are only efficient in decoding black and white visual tags.

Second, input from participants indicates that having a name (such as name of 2D visual tagging system provider) or some form of identifier on a visual tag enhances tag recognition and/or recall. This is probably based on the fact that at a psychological level, individuals can relate better to a named entity than a nameless brand. A named brand feels familiar, convenient and safe, while a nameless brand or technology is equated with inferior quality or status.

Third, it will be good practice to have some form of descriptive text beside a 2D visual tag. The descriptive text should describe what the visual tag stores and how users may interact with it. This practice (having descriptive text beside visual tags) will help novice users who are unfamiliar with tagging technology. It will also help those individuals who are already familiar with visual tagging systems as the descriptive text will help them to accurately and more quickly decode a visual tag. Although novice users may be drawn to a visual tag because of its distinctive design and name recognition, they may not know how to interact with the tag when they are near it, except information about the interaction procedure is provided.

V. FUTURE WORK

This study focused mainly on the evaluation of user perceptions about interacting with the BeeTagg, UpCode, ConnexTo, ShotCode and Semacode 2D visual tagging systems. But visual tagging technology is in a flux, and new visual tagging systems are increasingly being introduced into the market. There is therefore room for the extension of the scope of this study to other visual tagging platforms. It would also be expedient to carry out this type of study using a variety of camera phones.

ACKNOWLEDGEMENT

The author wishes to thank Dr. Audrey Mbogho for her supervision of this project.

REFERENCES

[1] Belt, Sara, Greenblatt, Dan, Häkkilä, Jonna, and Mäkelä, Kaj. UserPerceptions on Mobile Interaction with Visual and RFID Tags. Mobile HCI 2006.

[2] P. Diego L´opez de Ipi˜na, R. S. Mendonc¸a, A. Hopper. TRIP: A Low-cost Vision-based Location System for Ubiquitous Computing. Personal and Ubiquitous Computing Journal, Springer, 6, 3 (2002), 206–219.

[3] Johnson-Laird, P.N., Girotto, V., Legrenzi, P. Mental Models: A Gentle Guide for Outsiders. Sistemi Intelligenti, 9 (1999), 63-83. http://www.si.umich.edu/ICOS/gentleintro.html.

[4] Nielsen, Jakob. Usability Engineering. Academic Press, 1993.

[5] Rekimoto, J. and Ayatsuka, Y. CyberCode: Designing Augmented Reality Environments with Visual Tags. In Proc. DARE 2000 on Designing Augmented Reality Environments, ACM Press (2000), 1-10.

[6] Rohs, M. Visual Code Widgets for Marker-Based Interaction. Proc. 25th IEEE Int'l Conf. Distributed Computing Systems (ICDCS Workshops), Int'l Workshop Smart Appliances and Wearable Computing (IWSAWC), IEEE CS Press (2005), 506–513.

[7] Rohs, M. Real-World Interaction with Camera Phones. Proc. 2nd Int'l Symp. Ubiquitous Computing Systems (UCS 04), LNCS 3598, Springer-Verlag (2005), 74–89.

AUTHOR

Samuel Olugbenga King is with Mathematics Education Centre,Loughborough University, UK, E-mail: S.O.King@lboro.ac.uk

Manuscript received 8 November 2007. Published as submitted by the author.