



UNIVERSITY
of
GLASGOW

Brewster, S.A. and McGookin, D.K. and Miller, C.A. (2006) Olfoto: designing a smell-based interaction. In, *Conference on Human Factors in Computing Systems, 22-27 April 2006*, pages pp. 653-662, Montréal, Québec, Canada.

<http://eprints.gla.ac.uk/3252/>

Olfoto: Designing a Smell-Based Interaction

Stephen A. Brewster, David K. McGookin and Christopher A. Miller

Glasgow Interactive Systems Group, Department of Computing Science

University of Glasgow, Glasgow, G12 8QQ, UK

{stephen, mcgookdk}@dcs.gla.ac.uk www.dcs.gla.ac.uk/~stephen

ABSTRACT

We present a study into the use of smell for searching digital photo collections. Many people now have large photo libraries on their computers and effective search tools are needed. Smell has a strong link to memory and emotion so may be a good way to cue recall when searching. Our study compared text and smell based tagging. For the first stage we generated a set of smell and tag names from user descriptions of photos. Participants then used these to tag photos, returning two weeks later to answer questions on their photos. Results showed that participants could tag effectively with text labels, as this is a common and familiar task. Performance with smells was lower but participants performed significantly above chance, with some participants using smells well. This suggests that smell has potential. Results also showed that some smells were consistently identified and useful, but some were not and highlighted issues with smell delivery devices. We also discuss some practical issues of using smell for interaction.

Author Keywords

Olfaction, Smell, Digital Photographs, Searching, Tagging.

ACM Classification Keywords

H5.m. Information interfaces and presentation: Miscellaneous.

INTRODUCTION

Smell output is an under-explored presentation technique in human-computer interaction. Our sense of smell is very rich, but not often used. It is a powerful alerting mechanism and has strong links to both emotion and memory. These make it an interesting area of study for future user interfaces.

One reason for the lack of research is that there are few effective computer controlled smell devices which can emit a range of odours. Commercial devices are beginning to appear, but are not common. Many companies have proposed systems, but few have come to market. It is, however, still possible to investigate interactions using very simple

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2006, April 22–28, 2006, Montréal, Québec, Canada.

Copyright 2006 ACM 1-59593-178-3/06/0004...\$5.00.

smell technologies. Our aim in this work is to explore how smell might be used in a real application, if it is effective, and to discover the issues and practical problems that must be dealt with when designing smell-based interactions.

Our application focus is on searching digital photographs. Many people now have very large digital photograph collections with thousands of images. There are tools to support browsing and searching in many consumer-oriented applications (for example, Adobe Photoshop Elements – www.adobe.com, or Corel Photo Album – www.corel.com). These do not always provide the easiest ways to find photographs and users may not always use the tagging and searching features as they are too complex [11]. Our aim is to examine the use of smell and its link to memory and recall to aid photo searching by allowing users to tag photos with particular smells and then use those smells to help recall the photos later.

SMELL RESEARCH

The whole area of smell is much less well understood than vision or audition. We have no detailed understanding of how we perceive smell. We cannot easily build up complex smells from simpler components. This makes designing interfaces that use them more difficult, requiring rigorous evaluation to ensure usability.

As Kaye suggests in his thorough review of the whole topic [9], there are no good classification or description schemes for smells. There are specific ones for the perfume, wine and beer industries, for example, but these do not apply to the wide range of smells that we might want to use in a user interface. This makes it difficult to describe what is being used or to know what a smell you might want to use actually smells like.

Anosmia (the inability to smell one or more specific odours) also has a big impact. Large groups of the population are unable to smell particular smells with reasons for this only now being understood. Care must be taken to avoid these and evaluation is needed to ensure people can smell all of the smells you choose to use.

One of the key features of smell is its close link with memory. You may smell something and it takes you right back to the time and place that you smelled it before. Neuroscience research has shown strong links between smell and working memory, attention, reaction times, mood and emo-

tion [10]. If we can tap into these links then there is potential to use smell as an aid to searching or other interactions.

Cann and Ross [6] report that there is a strong link between smell and memory “*One experience related to the influence of odors that seems nearly universal, at least within the folklore surrounding odors, is the occurrence of vivid memories, often involving events from the distant past, precipitated by exposure to an odor associated with the original experience*”. They report a study by Laird in which 80% of men and 90% of women reported multiple odour-revived memories. The memories were often emotionally charged, rather than neutral. Cann and Ross say “... *it is clear that people can respond to olfactory cues as memory devices, describing events associated with odors, and that they believe odors often lead to the recall of past events*”. The kinds of memories evoked are often complex images or experiences rather than objects. This suggests that there is a strong response to smells that we could make use of for photo recall. It would seem essential to have a range of smells, both good and bad, to suit a wide emotional range.

Aggleton and Waskett [1] found a similar powerful effect with smells used in the Jorvik Viking Museum in York, UK with visitors recalling information about the museum significantly better in the presence of the Jorvik smells over six years later (smells at the museum are from Dale Air).

Cann and Ross [6] studied the use of odour to cue recognition of images. Participants were exposed to one of two odours when learning images and then asked to recognise them with and without the same odours present. They used a pleasant (flowery) or unpleasant (ammonium sulphide) smell during the learning phase. Results showed that recognition of the images 48 hours later was significantly better when the same smell was present as the learning phase.

The idea of such state-dependent recall is a useful one for our application as it gives us two choices of smells (abstract and representational) and two opportunities for associating a smell with a photo. Using representational smells that embody the original photo means that we can help people recall that original event. However, learning will also occur at the tagging stage where people associate smells with photos. Smells in this case could be either representational or more abstract (which might be easier to differentiate). Both types, when smelled again, would help recall of photos. This gives a real application the possibility of using a larger range of smells as we are just not tied to ones representing the actual contents of the photos, although the representational ones may be the most effective as they would be reinforced at the tagging stage, giving users more chance to associate them with the photo.

Cain [5] showed that certain smells appear to be more identifiable than others: coffee, peanut butter, Vicks, and chocolate were shown to be the most easy to identify, with cough syrup, cleaning fluid, and lighter fluid being the hardest. There is little practical work to help the interface designer choose smells. Much of the psychophysical work has used

very simple compounds, a problem then being naming the thing being smelled. For our purposes we wanted more ‘natural’ smells to represent the real-world things that people might have in their photographs.

There are few studies into the perception of intensity of odours. Engen and Pfaffman [7] showed that people could recognise three levels of the smells they used. This was less important for our study as the technology we used to deliver our smells had no control over intensity. With more sophisticated output devices one could perhaps use intensity to give some indication of the importance of the thing tagged, but this would be hard to control.

Smell Delivery

One important issue is smell delivery. As discussed, there is no simple way to mix complex smells from atomic ones, therefore most systems just use pre-packaged smells. The quality of the smells is often very variable; some smell very little like the thing they are supposed to, are quite synthetic smelling and hard to give a name to. This is a problem for interface designers as smells are required that people can identify. Kaye offers some help as he reports that “of personal memories evoked by an odor, 32% were evoked without an odor name”. This may mean that lower quality smells may still elicit the memories we need, even if they are not good enough for people to be able to name.

One key issue is that smells may diffuse across a room and thus be smelled by many people. This may be beneficial for an ambient display, but problematic for more specific ones. Yanagida *et al.* [16] looked at using an air cannon that tracks the user’s nose to present smells. This fires a focused vortex of air so that others nearby do not smell anything. Their initial prototype shows promise but is not commercially available. There are several examples of other devices that have been proposed but have never made it to market (DigiScents and TriSenx are well know examples).

For situations where there is just one user in an office or the smell can be displayed to a larger group of people things are easier. Simple fans can blow across the surface of scented material to spread it across a room, or the smell may just diffuse naturally.

For our study we used smell cubes from Dale Air, UK (www.daleair.com). These are 50mm³ green plastic cubes that contain an oil-based aroma on cotton wool. When the plastic tape covering the holes on top is removed, the smell diffuses out into the surrounding air. These are very cheap and are available in a wide range of different odours.

Uses of smell for interaction

There is little work in the area of smell and HCI. It has been used for ambient displays [13], but most uses have focused on notifications. Kaye has done some of the first work [8, 9]. He has proposed and built several systems that have investigated the use of olfaction for providing ambient notification information. His first system, *inStink*, provided a

sense of presence in a remote location. A spice rack with sensors attached to each spice jar was linked to a set of spray guns at a remote location. These were filled with essential oils which smelled like the spices in the spice rack. When a spice jar was removed from the rack, the corresponding essential oil was released at the remote location, allowing users there to be aware of activity at the location of the spice rack. Although Kaye did not perform a detailed evaluation, he noted that when the essential oils combined they produced an unpleasant smell. This further highlights that care must be taken when using multiple scents.

Kaye's second system, *Dollars & Scents*, presented information about the current value of the NASDAQ stock exchange using two scents. In this system a device was constructed using solenoids and perfume bottles that would emit a mint smell if the market was rising, or lemon if it was falling. Kaye reports no particular reason for the choice of these two scents other than they "leverage linguistic idioms as mnemonics". He also noted that the original system was to have used rose as the scent for a rising market, but was changed since one of the residents at MIT Media Lab (where the system was installed) was allergic to roses. Although Kaye reports no specific evaluation of *Dollars & Scents*, there was a generally positive response and the smells quickly were incorporated into the culture of the environment. For example people passing and saying "Smelling of mint today".

Scent Reminder used the same technology as *Dollars & Scents* but was integrated with Microsoft Outlook. Appointments could be set and scents released as reminders. Unfortunately, as with Kaye's other systems, there were no detailed evaluation so it is difficult to determine the actual usefulness or issues with the systems he proposes. However, his applications do show the potential that smell has to be a useful addition to computer interfaces.

Arroyo *et al.* [2] looked at interruption modalities (heat, light, sound, vibration and smell) to see which caused the greatest disruption to their ongoing work. For their smells they chose glue and soy sauce. No justification is given for the smells used. Their results were inconclusive; individual differences masked the effects of particular modalities. People's backgrounds made them respond very differently to the different modalities.

In a similar study Bodnar *et al.* [4] looked at auditory, smell and visual stimuli for interruptions. They used extract of cloves and artificial eucalyptus smells. These were chosen from an initial set of 10 in a pilot study. They could be easily differentiated at intervals of 30 seconds. Their results showed that olfactory cues were less effective than audio or visual ones, causing participants to reduce the number of questions completed during the study. Some interesting issues arose from participants' comments. Many felt that the smells used were too similar and thus easy to confuse. Smells lingering in the air and confusing participants was also a problem. As with Arroyo *et al.*, the background of

the participants played a role. An aromatherapist had no problem with the smell cues, whereas other participants said lack of experience with smell cues made the task difficult. These are serious issues for a photo searching application as users will have a wide range of backgrounds. One potential solution is to provide a wide range of different smells so that users can choose ones that suit them best.

Smell has been proposed in several VR systems to improve fidelity [15]. For example Tominaga *et al.* [14] used an air pump system to present aromas in virtual environments. They suggest two types of smells: scent of the environment and scent of the object. The environmental odours provide the smell of a place, to help the users feel like they are in a particular environment. The objects' odours give information about a particular object in the environment. Results showed very positive results with 5 aromas. People responded more positively about being in particular locations (forest, bar) when smells were present.

Tillotson from the Royal College of Art in the UK has investigated the use of smell for fashion, health and wellbeing, using smell emitted from clothing to create a 'scent bubble' around the wearer that "enhances the visual message of fashion with medical, sensory & psychological wellbeing for the wearer" (www.smartsecondskin.com).

PHOTOGRAPH CATALOGUING AND SEARCHING

Over recent years there has been a rapid growth in the use of digital cameras and many people have large numbers of photos on home computers which they need to search. There have been few published studies that have investigated the management of personal photos and people's cataloguing, searching and browsing habits. Rodden and Wood [11] undertook one of the most detailed studies to date using their ShoeBox digital photo cataloguing system with 13 users. They found a range of organizing strategies, from users putting images into different folders, to all images being in one large folder. Some users spent more time organising their photos, others spent very little time. Those who spent more time often changed the name of the folder to help them remember what was in it. Eight of their participants added extra text or spoken audio annotations to help searching. The spoken annotations were generally of people and places. One issue with the spoken annotations was the problem with speech recognition errors that could then make searching a problem. The methods for adding annotations or meta tags to facilitate searching need to be simple and effective otherwise people may not use them.

Rodden and Wood found users made three main types of queries:

- The set of photos from a particular event, e.g. holiday;
- An individual remembered photo;
- A set of photos taken at different events, but all sharing a property, such as containing a certain person.

Users in their study tended to do much more browsing than querying. One reason for this could have been that the users

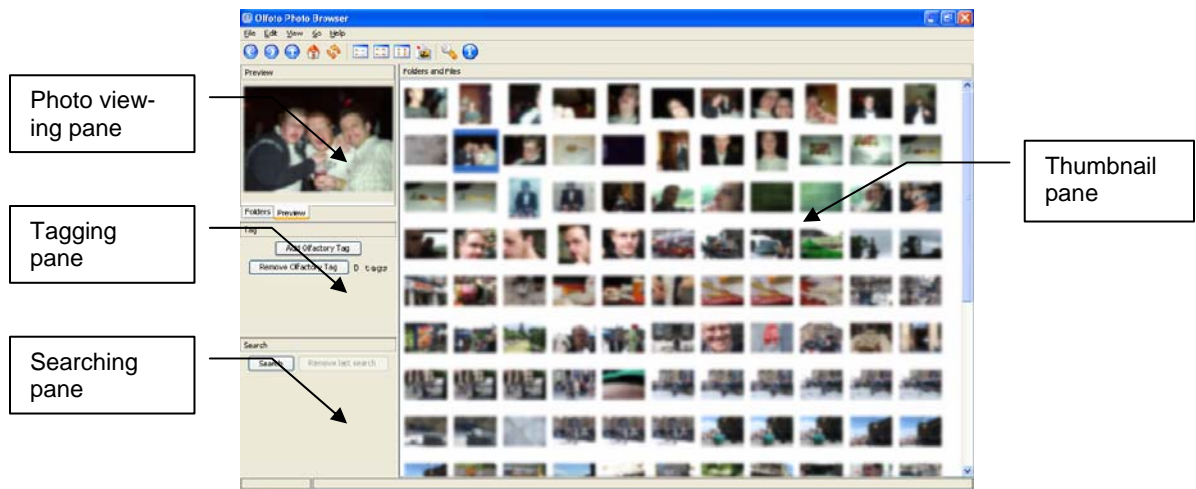


Figure 1. The Olfoto display.

were all novices with small collections of photos taken specifically for the study. As Rodden and Wood say “...queries might start to seem more important as a collection grows, and the photos get older and less familiar”. Participants in our study brought their existing photos with them to avoid this problem. We also wanted to make our tagging process as simple as possible, to avoid complexity that might put people off from using smell.

Shneiderman’s PhotoFinder [12] looked at simpler and more effective ways to tag photos, particularly people in photos. These were very low effort, simply requiring a user to drag a label onto a photo. These techniques are now common in commercial tools. We aimed to build on the simplicity of this approach in our photo tagging tool to make adding tags as simple as possible.

Bedersen’s PhotoMesa [3] provides sophisticated layout and visualisation mechanisms to help users find images from large sets of thumbnails. But with a large number of photos this can still be a problem. Alternative mechanisms to help cut down the number of thumbnails that needed to be presented would ease the problem.

We attempted to build on these previous systems, but use smell as a tool to tag and then retrieve photos. We also wanted to conduct an initial user study as many systems in this area have had little user evaluation.

OLFOTO

We developed a simple olfactory photo browsing and searching tool, Olfoto, based on the freeware Ekspos photo viewer (www.kiyut.com/products/ekspos). This provides basic thumbnail/photo viewing capabilities and, as it is written in Java, we could easily add our own code to allow smell and text tagging and searching (see Figure 1). It has a standard scrolling thumbnail pane occupying most of the screen, with a larger version of the chosen thumbnail shown top left in the photo viewing pane. Below this is the tagging pane and below that the search pane. Clicking on any of the

thumbnails in the thumbnail pane shows it in the photo viewing pane.

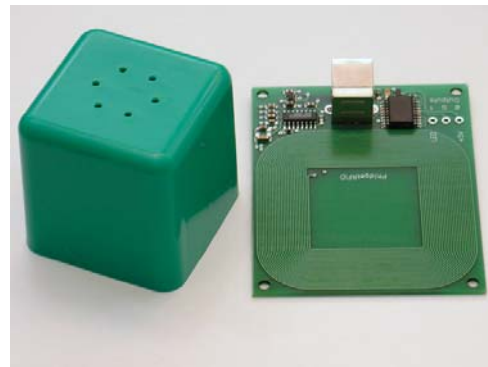


Figure 2. A smell cube and the RFID tag reader.

To present the odours we used smell cubes from Dale Air. An RFID disc tag was attached to the base of each of the cubes (Figure 2). We used Phidgets for this (www.phidgets.com) as they provide a very simple way of using RFID in Java. In tagging mode, users add one or more tags to a photo by moving the appropriate smell cube over the RFID reader. Multiple selections could be made in the thumbnail window so that the same smell could be added to multiple photos. Multiple smell tags could be added to a photo by moving more cubes over the reader. To remove a tag the cube was moved over the reader again. We tried to create a simple, tangible interaction that was easy to do, to avoid the problems highlighted by Rodden and Wood.

In search mode, the user moved a cube over the reader and all of the photos with that smell were displayed in the thumbnail pane. Multiple tags could be used for searching by passing the cubes of interest over the reader.

Text tagging and searching were done in a similar way; a user could choose one or more of the text tags from a drop down list and apply them to the selected photo(s). For searching the user chose the text tags from a drop down list to display the thumbnails with that tag attached.

EVALUATION

Our evaluation was done in three stages:

- *Categorisation Study*: Participants generate a list of smell and text tags to be used in the main studies;
- *Tagging Study*: Users tag photos with the smell and text tags from the first study;
- *Recall Study*: Users return after two weeks to search for photos with text and smell tags that they used in the second study.

The aim of the studies was to find out if smell could play a role in photo searching, how it compared to text tagging, if users would consider using it, and for us to understand how to use smell in a practical setting. This was our first study in this area and we needed to understand how to manage the practicalities of smell-based interaction.

Our study was undertaken in a usability lab which had two doors – one into the building and one to the outside – so we could get a good air flow through the room (to avoid the problems smell mixing found by Bodnar *et al.* [4]).

Our participants were all undergraduates, PhD students or members of staff in our Department. This was intentionally not a wide range. We needed to keep the likely range of tags small enough that we could cover them with a small selection of different smells to make the study feasible.

Categorisation Study

In this phase we wanted to generate a list of smell and text labels that could be used to tag photos for the main two studies. We used three participants, plus the three authors of the paper. Each was asked to bring along 300-500 photos from their collection. The experimenter then went through the photos asking each participant what smells labels he/she would use to describe the images; the process was then repeated for text labels. They were free to use whatever words came to mind and these could be different for the smell and text labels. We then grouped these words to come up with an overall set of smell and text categories. Table 1 shows the category names we chose based on the words generated by the participants for six of the smell categories.

We came up with 16 smell categories (Alcohol, Pine, Food, Fresh, Beach, Smoke, Garden, Musty, Grass, Floral, Sea, Sweat, River, Perfume, Petrol, Chocolate). These covered all of the words used, except for those that would have formed single word categories (i.e. words that could not be grouped with any others). We wanted to keep the number of categories small enough to make the rest of the study feasible. There were 12 text categories (City, Colour, Countryside, Family, Friends, Hobby, Holiday, Home, Party, Pub, Season, Sport, Travel, University, Work), again avoiding single word categories. These are by no means a complete classification of general photo searching keywords (which would be an interesting study). We used a focused group of users to make a set of tags that would likely be useful for a similar group of main test participants.

We then used the smell categories to buy a set of smell cubes from Dale Air. For this study we chose to use representational smells as a starting point given the research presented above.

Category name	Alcohol	Pine	Food	Fresh	Beach	Smoke
Words used	Wine x 3	Wood	Soup	Vanilla	Suntan lotion	Soot
	Vodka x 2	Washing up liquid	Pizza x 2	Snow x 2	Seaside	Smoke x 4
	Spirits	Tree x 3	Pasta sauce	Smirry rain	Sand x 4	Cigarette smoke
	Mulled wine	Sawdust x 2	Nuts	Ice cream	Beach	
	Morgans	Pine x 4	Garlic	Fresh water		
	Lager	Christmas Tree x 2	Fruit	Fresh smell x 2		
	Drink x 2	Alpine	Food x 4	Fresh morning		
	Cider		Curry	Dew		
	Beer x 2		Burgers	Air con x 2		
	Alcohol x 2					

Table 1. Six categories of smells with the number of occurrences of each of the words used.

A major problem was taking the smell categories and matching them to a smell produced by Dale Air. The company produces around 200 different odours, and without smelling them, it was hard to know which were the ones most useful to us. This is an example of the problem that Kaye identified above – there is no standard classification scheme for smell. A floral smell from one company may be completely different to that of another (and still not like a real floral smell). In addition, some of the names are not very informative (Dragons breath, for example); it is difficult to know what the smell is like and if it matched the category we had, but we did our best. The smells chosen from the Dale Air website to match our categories were: Brewery, Alpine, Bread, Ozone, Sea Shore, Smoke, Farmyard, Dusty, Grass, Floral, Sea Breeze, Sweaty Feet, Riverbank, Unisex Perfume, Machine Oil, Dark Chocolate.

Tagging Study

The next stage of the work involved participants tagging photographs with the smell and text labels generated in the previous stage. We used 12 new participants who were paid £10 if they completed both this and the following Recall Study. There were four women and eight men, ages ranging from 20-45, six were from the UK, two were Indian, one each were from China, Mexico, Greece and South Africa. All were asked to bring along around 500 digital photos from their collections. This gave us a wide enough range of material for them to be able to tag and also to create a large enough set for the Recall Study.

We asked what software they used to manage their photo collections. Six did nothing, just putting them into a folder on their PC, two used the online photo service Flickr (www.flickr.com), two used iPhoto and one used the Canon software that came with his camera. We asked if they

Smell Name	Floral	Ozone	Riverbank	Dusty	Alpine	Unisex Perfume	Grass	Sweaty Feet	Dark Chocolate	Farmyard
Participant labels	Sweet Sawdust	Tree Bark	Musk	Engine	Perfume	Mild Perfume	Sickly	Pale oil	Chocolate	Horse
	Soap	Deodorant	Olbas Oil	Sport shop	Talcum	Orange	Grass	Sweat	Chocolate	Sh*te
	Soap	Lily	Mint	Apple	Lemon	Orange	Wood	Cider	Chocolate	Rubber
	Bleach	Toothpaste	Vicks	-	Perfume	-	Flower	Cheese	Caramel	-
	Lavender	Menthol	Menthol	Painted Wood	Talcum Powder	Shampoo	Cut Grass	Trash Can	Coffee Chocolate	Wet Board Cleaner
	Talcum Powder	Soap	Herb (Oregano)	Dust	Air freshener	-	-	Sweaty feet	Chocolate	Sh*te
	Trees	Water	Menthol	Smoke	Perfume	Orange	Grass	Sweaty feet	Chocolate	Sh*te
	Bathroom Spray	Flowers	Vicks	Toilet Cleaner	-	Perfume	Pepper	Feet	Chocolate	Incense
	Soap	Mint	Vicks	Campfire	Soap	Bathroom Cleaner	Grass	Cheesy Feet	Chocolate	Poo
	Talcum Powder	-	Sweet Pepper	Wet Wood	Strong Flowers	Sports Building	Leaves	Compost	Cake	Bad Smelling Compost
Culicura	Vicks 2	Vicks	Bad Smell	Powder	-	Fruit	Pajhem	Raspberry	Bad Smell	

Table 2: Names used by participants to describe smells in the Tagging Study.

tagged their photos in any way. Five did nothing, four renamed files with suitable labels, one Flickr user and one iPhoto user used text annotations and one iPhoto user created albums, used text annotations, created titles and keywords for his photos. This is generally in line with the results of Rodden and Wood [11].

Participants either did the text tagging first or worked through the smell cubes and assigned a name to each one and then did the smell tagging (in a counterbalanced order). Each tagging phase took around 20 minutes. Finally, we interviewed the participants about the process of tagging and general issues raised. The whole process took one hour.

When tagging with the smells participants were free to pick up any cube they desired and smell it by lifting the tape. The cubes were not labelled in any way. The cubes were on the table next to the PC running our software and they could smell (and tag with) as many different cubes as they wanted. If they were happy with a chosen smell they could move it over the RFID tag reader to tag the photo.

Results and Discussion

Table 2 shows the names participants gave to 10 of the 16 smells. The first five were less well identified; the last five were much more consistently named. One problem was that some of the smells were very synthetic smelling. Floral, for example, had a very soapy, bathroom cleaning product smell and many participants identified it as that rather than flowers. As suggested by Cain [5] chocolate was well recognised. His list also included Vicks (a gel rubbed on the chest to ease a blocked nose). It is interesting to note that four of the participants identified Riverbank as having that smell. Again this was due to its very synthetic nature.

These results show that some smells could work really well, some do not smell like their name but are consistently identified by participants, and some are not consistently identified at all. It is useful to know consistency for providing a

general set of smells for an olfactory photo tool. It means that a set of easily differentiable smells could be provided with the system. However, a more important issue is how consistently people can name the smells over time, which was investigated in the Recall Study below.

Further work is needed to build up a set of good smells. The only way this can be done is empirically because, as Kaye reports, there is very little research to help choose a wide range of uniquely identifiable smells.

Participants completed NASA TLX subjective workload scales after both conditions. These give subjective data on how demanding people feel a task to be and are a good complement to quantitative measures of error rates, giving a more complete picture of usability. We added a question on fatigue (to see how tiring the task was) and on which method was preferred (participant 6 failed to complete the smell part of the subjective scales, so has been excluded from the analysis). Figure 3 shows the results. An analysis of Overall Workload (using, as for all the analyses in the paper, a paired two-tailed T-test) showed a significant reduction for the Text condition (Smell=10.4, Text=6.8, $T_{10}=3.84$, $p=0.001$, $\eta^2 = 0.24$). We then performed a detailed analysis of each of the individual factors. Results showed a significant reduction in workload in the Text Condition for: Mental demand ($p=0.005$), Physical demand ($p=0.006$), Perceived performance ($p<0.001$) and Fatigue ($p=0.01$). Temporal demand, Effort expended, Frustration and Preference showed no effect.

These results show that it was easier for people to tag photos with text labels than with the smells. Participants found the task mentally demanding as they were not used to using smells in this way and it required some concentration. This was reflected in the physical demand, fatigue and the low level of perceived performance. It was interesting to note that there was no effect for frustration or preference given the above results. This suggests that people might persevere

Smell Name	Floral	Ozone	Riverbank	Dusty	Alpine	Unisex Perfume	Grass	Sweaty Feet	Dark Chocolate	Farmyard
Participant labels	Perfume	Faint Linseed	Pale Musk	Tack Room	Sickly Sweet	Faint Perfume	Sweet Scent	Treacy Linseed	Chocolate	Horse
	Alcohol	Sugar Candy	Mint	Wood	Lemon	Apple	Grass	Rotten Egg	Chocolate	Rubber
	Perfume	-	Menthol	-	-	-	-	-	Caramel	-
	Mild Fragrance	Mild Frosty	Sweet Pepper	Slightly Mouldy Plants	Strong Talcum Powder	Talcum Powder	Fragrant Flower	Decaying Plant	Sweet Food	Mouldy Flowers
	Yardly	Fruits	Vicks	-	Tamarind	Medicine	Medicine	Fruit	Fruits	Store

Table 3: Words used by participants to describe smells in the Recall Study (gaps are where users made no response).

with the smells, and with some practice, they would become easier to use and less demanding.

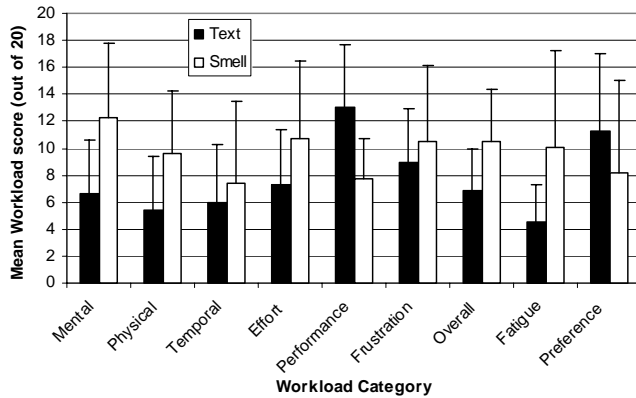


Figure 3: Means and standard deviations of subjective measures in the Tagging Study.

One of the most raised issues in the informal comments from participants was that the smells did not always match their photos. Many wanted to choose their own smells, which was not practical for this study. Some also said that certain of the smells were hard to tell apart (Ozone, Sea shore and Sea breeze were most commonly mentioned). This was a problem as we could not smell the cubes before we bought them. Participants also wanted to be able to create their own text tags (which all real photo-cataloguing systems do).

Recall Study

Two weeks after the Tagging Study participants returned to perform the second part of the task to see if smell helped them recall photos, and how this compared to the use of text labels. Only six of the original twelve were able to return for this second phase due to other commitments.

Participants performed the two conditions in the same order as the previous study. The first step got them to go through the smell cubes again and re-label them. This allowed us to get an idea of the consistency of their naming. They then answered three types of questions in both conditions. The range of questions tried to gauge their ability to use smell and text labels in a range of different ways and were motivated by the queries identified by Rodden and Wood [11].

1. *Smell/text* \Rightarrow *photo*: Participants were shown 4 photos and given one smell or text tag. They had to identify which photo had that tag. Only one photo in the set of four had the tag and the three distracters were chosen randomly from the rest of their photo collection. They had to do this 15 times with different photos. We measured the number of correctly identified photos.
2. *Photo* \Rightarrow *smell/text*: Participants were given four smell/text tags and one photo. They had to identify which tag was attached to the photo. This was done 15 times. We measured the number of correctly identified tags.
3. *Search*: Participants then performed a more general search task. They were given three key features of a photo and then asked to find it in their photo catalogue using smell or text tags. We chose the photos at random and picked the key features ourselves. For example, we might ask for photo with a building, people and palm trees, or an arch, angels and a roof. They did this five times. In the Smell Condition, the cubes were on the table next to the PC and participants were free to smell them as they liked. If they were happy with a chosen smell they could move it over the RFID tag reader to search with it. We measured the number of tags that they tried before finding the right photo. The fewer tags used would suggest that that type of tag helped them find the particular photo better. It was possible for a number of photos to match the three criteria we gave the participants. If the photo identified matched the criteria, they were given a mark.

They were presented with similar questions for both conditions. Participants filled in subjective workload assessment sheets as before after each condition. Finally, we interviewed them about the study as a whole to get their views. The study took around 90 minutes.

Results and Discussion

Table 3 shows the same example smells as Table 2. Chocolate was still well identified in this second phase. Sweaty feet and farmyard were much less well identified. Floral still had a very perfumed smell that was commonly identified by participants as was the Vicks/minty smell of Riverbank.

These results show that there is some similarity in the judgements made by the participants, but there is less consensus in this second phase (partly due to the smaller number of participants). It is not clear why they would be less good in this second phase, except if the smells were beginning to fade. By the time the second study took place, the cubes had been open for a month and a half. Some of the strong initial smells had definitely mellowed and changed. This makes it hard to assess our participants' abilities to re-identify the smells as they may not have been the same. For future work it would be worth using a set of new cubes for the second study to see if the effect disappeared. This also had an impact on their answers to the questions we set in this phase. If the smells were different then people were likely to do worse as the correct memories would not be evoked. There was no variation in the text tags, so performance should be better.

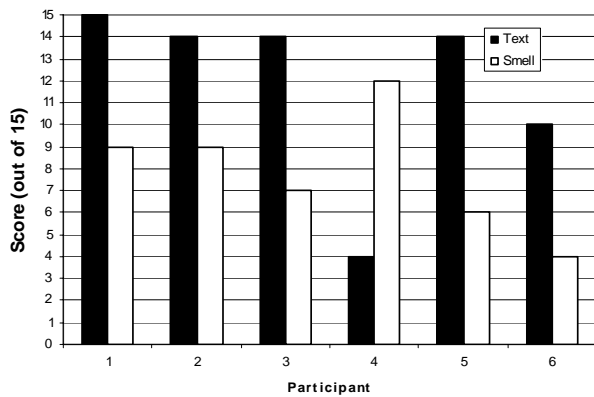


Figure 4. Number of correct answers for question type 1 in the Recall Study.

The results for question type 1 showed no significant difference between conditions (mean correct score of 7.8 (out of 15) with Smell and 11.8 with Text, $T_5=1.64$, $p=0.16$, $\text{Eta}^2 = 0.27$). Formal statistical analysis is limited due to the small number of participants. Looking at the data in Figure 4 a trend towards text tagging can be seen, with participant 4 behaving somewhat differently to the rest. This suggests we too were seeing the marked individual differences that Arroyo *et al.* [2] found. However, participants were right using the smells 52% of the time, chance would be 25%. This shows that they were able to use the smell tags. With new cubes and better smells we might improve upon this performance.

The results for question type 2 showed the text tags being significantly better than the smell ones (mean number of correct identifications was 8 for Smell and 14.5 for Text, $T_5=3.4$, $p=0.01$, $\text{Eta}^2 = 0.53$). Participants were very good at matching a text label to a photo. Performance in the Smell Condition was at the same level as for question type 1, with participants right 53% of the time. Results are shown in Figure 5. Participant two was particularly bad in the Smell Condition, whereas she had performed well with question

type 1. Participant four again performed reasonably well with this question type.

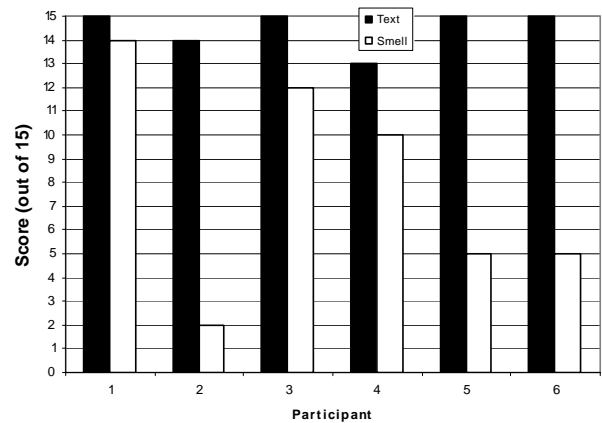


Figure 5. Number of correct answers for question type 2 in the Recall Study.

The results for question type 3 (Figure 6) showed no significant difference between smell and text tagging. All users used the same or fewer tags to find the photos in the Text Condition (mean of 3.6 tags used in the Smell Condition and 1.5 in the Text, $T_5=2.13$, $p=0.08$, $\text{Eta}^2 = 0.27$). Participants again showed marked individual differences, with some participants using many more tags in the Smell Condition, and others using the same number as for Text. We need a study with a larger number of participants to understand these effects fully.

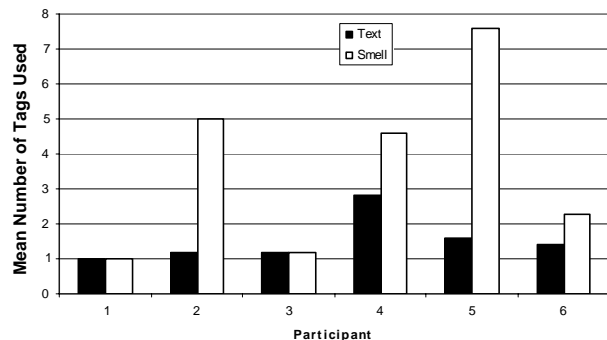


Figure 6. Mean number of tags used for question type 3 in the Recall Study.

The results of the subjective measures are shown in Figure 7. Participant 4 failed to complete one part of the subjective scales, so was removed from the analysis. Analysis showed no effect for overall workload (Smell=12.3, Text=5.9, $T_4=2.5$, $p=0.06$, $\text{Eta}^2 = 0.46$), so no further analysis was undertaken. However, looking at the figure there seems to be a clear trend in favour of the Text Condition (with all categories in the Text Condition showing a reduction) and the pattern of the data is very similar to that of the Tagging Study, with perhaps a larger difference between the perceived performance with smell and text tagging. With a larger number of participants it is likely the differences would be significant. There were no significant differences

for Fatigue or Preference. The scores for fatigue were almost identical to the Tagging Study, preference for the text tagging had increased a little from the earlier study.

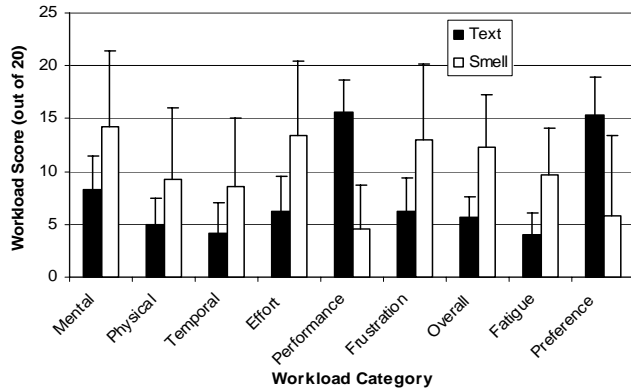


Figure 7. Means and standard deviations of subjective measures in the Recall Study.

From informal comments many participants felt that the whole interaction with the smells was good and easy to use. However, two people said that they felt they were trying to remember the smells, rather than being ‘taken back’ to the photos they were looking for. This may have been because the smells changed slightly between the two stages of the Study. Differentiability of the smells was also mentioned again – people again found it difficult to tell some apart, so this affected their searching.

DISCUSSION

In general, participants were able to use the smells to identify pictures. Participants were performing at greater than chance levels with the smells. They were less effective than text tags, but this is perhaps not surprising. Using smells in this way is an unusual thing for people to do, whereas text labels are much more familiar. With greater usage performance is likely to improve. An added difficulty was that some of the smells may have changed since the participants learned them, reducing their power for recall. Even with these problems, participants were able to use the smells.

Is smell tagging really practical? One key issue would be getting hold of the smells you wanted to tag with. We provided a small set to our group of users and these did not cover all the things they wanted to tag. There would need to be some simple way to buy appropriate smells, perhaps in cartridges like printer ink, dispensed by machines located by digital photo kiosks (this was the method suggested by DigiScents for their device, but it never came to market). Lower odour intensity output devices are also needed, suitable for a single user type environment. Dale Air recently released their Vortex USB device (www.daleair.com) that uses much less powerful smell discs. They do not saturate the environment in the same way as the cubes do, so are much more suitable for single user environments. We are using these for our next set of studies.

We tried to allow users to make a natural mapping between photo and tag. This is difficult as a large number of different smells would be required. An alternative would be to use an abstract or semi-abstract mapping. One might tag photos of the beach with chocolate, which has no relevance to the scene, but is easily identified. This would remove the emotional link but might still help recall, as Cann and Ross showed [6]. We found this behaviour as one participant informally reported that he was using ‘nice’ smells to tag ‘nice’ photos and ‘bad’ smells for ‘bad’ photos when no suitable representational one was available.

For a further study it would be useful to try and get participants to describe the types of photos a smell brings to mind in the Recall Study, rather than just doing searching with the tags. This would give some more general, subjective information about the effect of the smells.

Much work needs to go into the choice of smells. As Kaye suggests [9], there is little practical research of use to interaction designers that can help in the choice. Some of the smells we ordered smelled quite similar (Sea Shore and Sea Breeze for example), others smelled little like we imagined when we ordered them. There are some (chocolate for example) that really worked well; we need further research to identify more of this type.

Some Practical Considerations When Using Smell

There turned out to be many practical issues to deal with when running our study. Some have been mentioned before and some occurred because of our design. We took care to make sure the room was well ventilated so that smells did not linger. Some participants did complain that the smells got mixed with each other. Due to our task users needed to switch between smells frequently. Even if the top was only off a cube for a short while it did get out into the room. The use of a fan might alleviate this problem, as would the use of less powerful smell discs.

When we first received the sixteen cubes the smells were very powerful and created a nasty mixture (which infused our whole department). We put the cubes in zip-loc bags and then in a sealed plastic box to stop the smells escaping. In the future we may leave new cubes open outside for a few days to let them ‘cool off’ before using them so that we can get an even level of intensity for our studies. After a week or so, the smells had evened out so that we could store them with no problems. Some of the smells were of very different intensity levels. Sweaty Feet was initially very strong, masking the other smells in early prototypes.

Smells rapidly diffuse across wide areas and can reach people for whom they were not intended. Keeping them under control is difficult. This affects the uses to which they can be put. The cubes we were using were not designed for single users, they were meant for larger rooms, shops or museum exhibits so had to be quite intense. For a real photo searching application much less powerful smells would be more useful. Since our study was completed, the Vortex

device has become available and solves many such problems.

After some pilot studies we took care to wipe the sides of the cubes. Pilot users had complained of getting smells on their fingers so that one smell then interfered with the next. At the start of each day we wiped the cubes so that there was no residue on them.

Related to the above is the important issue of consistency of the smell delivery device. Our smell cubes were perhaps starting to run out by the second study which meant that the smells had changed (some of the more volatile components may have evaporated, making the odours smell different). A computer-controlled device that kept the smells sealed and only opened them when needed might solve this problem.

Two participants said their sense of smell was impaired (and did not return for the Recall study). One had hay fever and one a cold. This is a particular issue for our olfactory sense as it is more susceptible to day-to-day variation than our sight or hearing. A good screening program for experimental participants might avoid this problem, but it would not deal with the issues of using a smell-based interface as part of a normal interaction with a computer. As with all such issues, a good multimodal system that used several different senses would be best. Users could then pick the interaction techniques most suitable; if they had a cold they could use text or audio searches instead that day. We will investigate the use of smell in a multimodal context in our next study.

CONCLUSIONS

The studies presented here have shown that smell can be used to aid photo recall. Participants performed significantly better than chance when using the smells, and several participants used the same number of smell and text tags to find the photos for which they were searching. Text tagging did work better overall, but the results were very positive for the first smell-based study of this type.

Many issues remain about the use of smell in HCI. Choosing an appropriate set of smells is difficult and more research needs to be done into creating an easily differentiable set. Individual differences also seem to play a large role; some people are better at using smells than others. A bigger study is needed to investigate this further. If some of these issues can be addressed, the research presented here shows that smell has a valuable role alongside the other senses in future user interfaces.

ACKNOWLEDGMENTS

This project was funded by the Department of Computing Science and EPSRC Advanced Research Fellowship GR/S53244. Thanks go to Dale Air for their help with the smell cubes.

REFERENCES

1. Aggleton, J.P. and Waskett, L. The ability of odours to serve as state-dependent cues for real-world memories:

- Can Viking smells aid the recall of Viking experience? *British Journal of Psychology*, 90 (1999), 1-7.
2. Arroyo, E., Selker, T. and Stouffs, A., Interruptions as multimodal outputs: which are the less disruptive? In *Proceedings of IEEE International Conference on Multimodal Interfaces (ICMI'02)*, IEEE Press (2002), 479-482.
3. Bedersen, B., PhotoMesa: a zoomable image browser using quantum treemaps and bubblemaps. In *Proceedings of ACM UIST 2001*, (Orlando, FL, USA), ACM Press (2001), 71-80.
4. Bodnar, A., Corbett, R. and Nekrasovski, D., AROMA: ambient awareness through olfaction in a messaging application. In *Proceedings of ACM International Conference on Multimodal Interfaces (ICMI'04)*, (State College, PA, USA), ACM Press (2004), 183-190.
5. Cain, W.S. What we remember about odors. *Perfumer and Flavorist*, 9 (1984), 17-21.
6. Cann, A. and Ross, D. Olfactory stimuli as context cues in human memory. *American Journal of Psychology*, 102 (1) (1989), 91-102.
7. Engen, T. and Pfaffman, C. Absolute judgements of odor intensity. *Journal of Experimental Psychology*, 58 (1) (1959), 23-26.
8. Kaye, J. Aromatic Output for HCI. *Interactions*, 11 (1) (2004), 49-61.
9. Kaye, J. *Symbolic Olfactory Display*. MIT, Media, Arts and Science, School of Architecture and Planning, Masters Thesis, 2001.
10. Michael, G.A., Jacquot, L., Millot, J.-L. and Brand, G. Ambient odors modulate visual attentional capture. *Neuroscience Letters* (352) (2003), 221-225.
11. Rodden, K. and Wood, K.R., How Do People Manage Their Digital Photographs? In *Proceedings of ACM CHI 2003*, (Fort Lauderdale, FL, USA), ACM Press Addison Wesley (2003), 409-416.
12. Shneiderman, B. and Kang, H., Direct annotation: A drag-and-drop strategy for labelling photos. In *Proceedings of IEEE Information Visualisation 2000*, IEEE Press (2000), 88-95.
13. Strong, R. and Gaver, B., Feather, Scent and Shaker: Supporting Simple Intimacy. In *Proceedings of CSCW '96*, (Boston, MA, USA), ACM Press (1996), 29-30.
14. Tominaga, K., Honda, S., Ohsawa, T., Shigeno, H., Okada, K. and Matsushita, Y., "Friend Park"-expression of the wind and the scent on virtual space. In *Proceedings of Virtual Systems and Multimedia '01*, IEEE (2001), 507-515.
15. Washburn, D. and Jones, L.M. Could olfactory displays improve data visualization? *Computing in Science & Engineering*, 6 (6) (2004), 80-83.
16. Yanagida, Y., Noma, H., Tetsutani, N. and Tomono, A., An Unencumbering, Localized Olfactory Display. In *Proceedings of ACM CHI 2003*, (Fort Lauderdale, FL, USA), ACM Press (2003), 988-989.