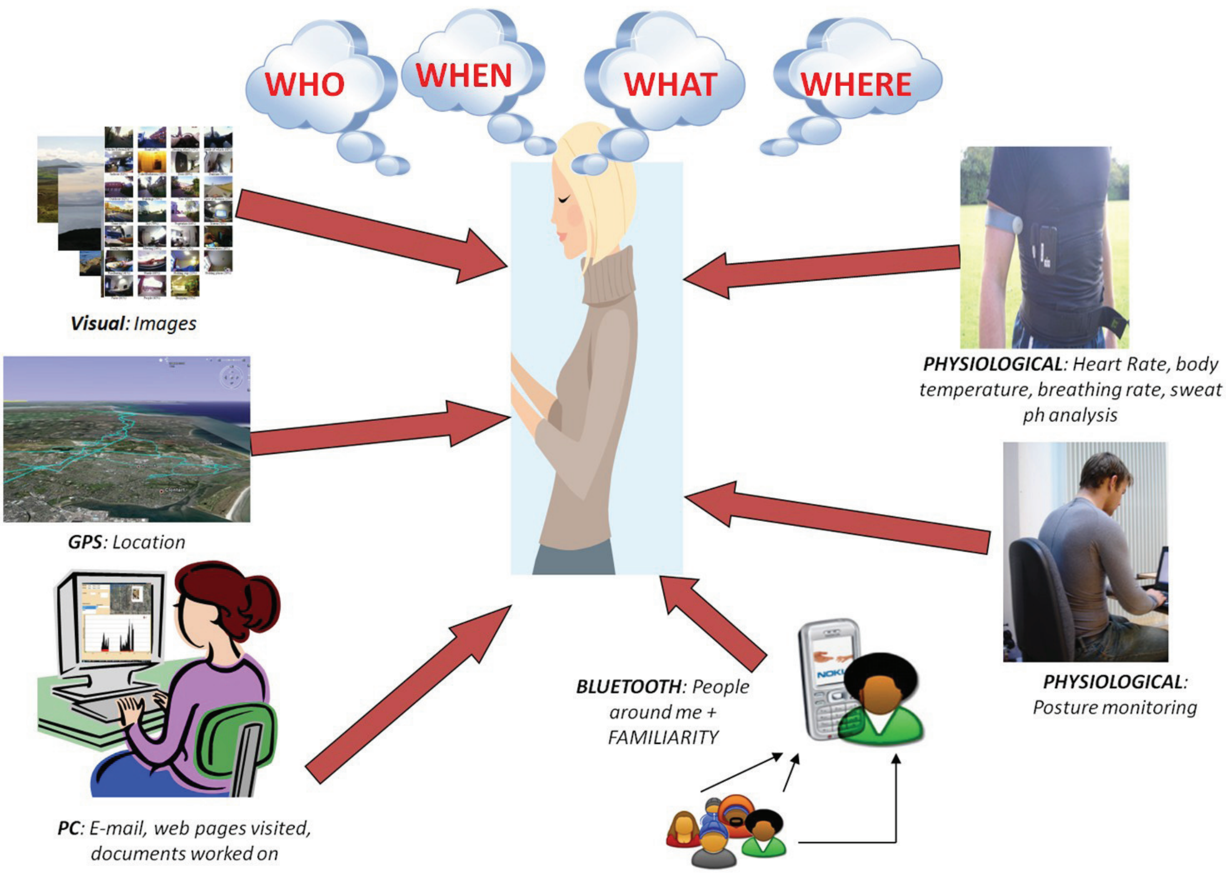
Information Access for Personal Media Archives

At ECIR 2010, 28th March 2010

CLARITY

Towards e-Memories: challenges of capturing, summarising, presenting, understanding, using, and retrieving relevant information from heterogeneous data contained in personal media archives.



www.iapma2010.dcu.ie

Information Access for Personal Media Archives

At ECIR 2010, 28th March 2010 CLARITY

Towards e-Memories: challenges of capturing, summarising, presenting, understanding, using, and retrieving relevant information from heterogeneous data contained in personal media archives.

Welcome to the inaugural workshop on "Information Access for Personal Media Archives". It is now possible to archive much of our life experiences in digital form using a variety of sources, e.g. blogs written, tweets made, social network status updates, photographs taken, videos seen, music heard, physiological monitoring, locations visited and environmentally sensed data of those places, details of people met, etc. Information can be captured from a myriad of personal information devices including desktop computers, PDAs, digital cameras, video and audio recorders, and various sensors, including GPS, Bluetooth, and biometric devices.

In this workshop research from diverse disciplines will be presented on how we can advance towards the goal of effective capture, retrieval and exploration of e-memories. We were delighted to have a keynote which was given by Dr. Daniela Petrelli who discussed the merits of digital technology and personal media archives, and the value in assisting people in effectively accessing their memories. CLARITY researcher Edmond Mitchell presented developments in the field of wearable sensing which will provide less intrusive monitoring of person activities in the near future (see page 4). Prof. Kiyoharu Aizawa from Tokyo presented an overview of one of the most successful systems in the lifelogging domain, namely a diet monitoring application which has approximately 280 users, who have uploaded more than 5,500 images of meals (see page 11). Kan Ren from the University of Surrey introduced a novel intuitive interactive interface for browsing of large-scale personal image and video collections (see page 14). Daniel Blank from the University of Bamberg in Germany considered the problem of managing the vast collections of lifelog generated content, and suggested that peer-to-peer (P2P technologies offer interesting solutions (see page 22). Ivar Solheim from the Norwegian Computing Center discussed the perceived and actual effectiveness of various search and navigation strategies used by individuals when interacting with their lifelog content (see page 30).

Organising Committee

Aiden R. Doherty (Dublin City University, Ireland)

Cathal Gurrin (DCU)

Gareth J.F. Jones (DCU)

Alan F. Smeaton (DCU)

www.iampa2010.dcu.ie **Published at http://doras.dcu.ie/15373/** **ISBN: 1872327869**

Workshop Program

Location: KMi Podium, Berrill Building, Milton Keynes, U.K.

Opening Session	**1.45 – 3.30 pm**
	Opening Remarks
Keynote by Daniela Petrelli	*Digital Technology and the Value of Personal Memories*
Shirley Coyle, **Edmond Mitchell,** Tomas Ward, Gregory May, Noel O'Connor and Dermot Diamond	*Textile Sensors for Personalized Feedback (Page 4)*
Kiyoharu Aizawa, Gamhewage Chaminda de Silva, Keigo Kitamura and Yuto Maruyama.	*FoodLog: the Easiest Way to Capture and Archive What We Eat (Page 11)*
Afternoon Coffee/Tea Break	**3.30 – 3.50 pm**
Session 2	**3.50 pm – 5.30 pm**
Kan Ren, Risto Sarvas and Janko Calic	*Evaluating Interactive Access to Personal Visual Archives (Page 14)*
Daniel Blank and Andreas Henrich	*Description and Selection of Media Archives for Geographic Nearest Neighbour Queries in P2P Network (Page 22)*
Ivar Solheim, Øystein Dale, Lothar Fritsch, Till Halbach, Knut Holmqvist and Ingvar Tjøstheim.	*Search and navigation as retrieval strategies in large photo collections (Page 30)*
	Closing Remarks

Program Committee

Kiyoharu Aizawa (Toyko, Japan)

Lothar Fritsch (Norwegian Computing Center)

Josef Hallberg (Luleå University of Technology)

Steve Hodges (Microsoft Research Cambridge)

Kieron O'Hara (University of Southampton, UK)

Mike Sinclair (Microsoft Research, USA)

Steve Whittaker (IBM Almaden, USA)

Kevin Curran (University of Ulster, N. Ireland)

Jim Gemmell (Microsoft Research, USA)

Alexander Hauptmann (CMU, USA)

Kälervo Jarvelin (University of Tampere, Finland)

Daniela Petrelli (University of Sheffield, UK)

Kristin Tolle (Microsoft Research, USA)

Textile Sensors for Personalized Feedback

Shirley Coyle
National Centre for Sensor Research, CLARITY Centre for Sensor Web Technologies, Dublin City University, Glasnevin, Dublin 9, Ireland
shirley.coyle@dcu.ie

Edmond Mitchell
Centre for Digital Video Processing, CLARITY Centre for Sensor Web Technologies, Dublin City University, Glasnevin, Dublin 9, Ireland
edmondmitchell@gmail.com

Tomas E. Ward
Biomedical Engineering Group National University of Ireland Maynooth, Maynooth, Co. Kildare, Ireland
tomas.ward@eeng.nuim.ie

Gregory May
Dept. of Human Health and Performance, CLARITY Centre for Sensor Web Technologies, Dublin City University, Glasnevin, Dublin 9, Ireland
gregory.c.may@gmail.com

Noel E. O'Connor
Centre for Digital Video Processing, CLARITY Centre for Sensor Web Technologies, Dublin City University, Glasnevin, Dublin 9, Ireland
noel.oconnor@dcu.ie

Dermot Diamond
National Centre for Sensor Research, CLARITY Centre for Sensor Web Technologies, Dublin City University, Glasnevin, Dublin 9, Ireland
dermot.diamond@dcu.ie

ABSTRACT

Wearable sensors provide a means of continuously monitoring a person in a natural setting. These sensors can "look in" by monitoring the wearer's health through physiological measurements and also by detecting their activities. Other sensors can be used to "look out" from the wearer into the environment through which he/she is moving, which may serve to detect any potential hazards or provide contextual information about the wearer's lifestyle. Wearable sensors can be harnessed to give immediate feedback to the wearer while also providing an archive of physiological data which can be logged and assessed over days, months or years. This has many applications in the field of healthcare, rehabilitation and sports performance. Here we present a number of case studies involving "smart" garments which have been developed to monitor the well-being of the wearer and to assess performance and progress, for example in training or rehabilitation scenarios.

Categories and Subject Descriptors

A.0 General

General Terms

Measurement, Performance, Design, Experimentation, Human Factors,

Keywords

Wearable sensors, physiological monitoring, body sensor networks

1. INTRODUCTION

Wearable sensors integrated into garments are adding new functionality to today's fashion. Smart garments have the ability to sense the wearer and also the environment surrounding the wearer. The advantage of using textiles as a substrate for sensor integration is that they are ubiquitous. Textiles used in garments we wear are in intimate contact with the body, and also form the interface between us and our environment. In addition, textiles exist in our surroundings e.g. home textiles such as curtains and upholstery. Therefore by developing textile based sensors there are many parameters that can be measured relating to the person and their environment. Wearable sensors to monitor a person's health give rise to a concept of personal health or pHealth. This concept means that the person can be more aware of their own well-being and can take a more active role in their health by managing their lifestyle to prevent illness. The pHealth concept emerged as a response to the unsustainable increase in healthcare costs world-wide due mainly to an epidemic of 'life-style' diseases arising from unhealthy diet and lack of exercise, such as obesity, cardiovascular disease, diabetes and chronic respiratory disease. The current global demographics show an increase in the aging population which also puts further burden on the healthcare sector. Wearable sensors may be used as a tool to gather information about the person's health and lifestyle to give a more complete picture. By continuously monitoring the wearer over time an archive of their personal health can be created which is far more desirable than the snapshot images that are currently attained with occasional clinical visits.
Sensors "looking in" to the body may be used to monitor daily activity. A simple approach is a pedometer to count the number of steps that a person has taken in a day. This may be used as a crude indicator of their level of physical activity. More sophisticated sensors can measure specific joint movements, position and speed that can be used to develop a model of kinematic movements. This may be used to assess what type of movements the person is undertaking and how well the person is performing an exercise. Sensors that measure body movements can be used for home rehabilitation based on prescribed physiotherapy exercises.

Physiological measurements are also possible with wearable sensors. Textile electrodes to measure electrocardiographs have been developed and shown to perform well compared to gold standard methods [1]. Shirts capable of monitoring vital signs have recently appeared on the market such as the Vivometics Lifeshirt [2]. Sensors pertaining to physiological measurement may be used to assess chronic illnesses such as cardiovascular disease. In others they may be used to help early detection and prevention of illness where conventional clinical visits can only provide a brief window on the physiology of the patient. Exercise is of huge importance in underpinning a healthy lifestyle, and providing incentives for people to exercise is essential for maintenance of personal health. The sports industry has identified a huge emerging market for wearable devices. The world of professional sports and athletics has seen dramatic changes in performance largely due to physiological testing and a better understanding of the effects of different training techniques on the body. Physiological testing ensures athletes stay healthy and develop personalized training strategies to keep the fine balance between over-training and reaching peak performance. At present, the vast majority of this testing is done in laboratory settings, yet wearable sensors now have the potential to allow physiological testing to be carried out in the natural setting, on the track, pitch or court. However, it is not just elite athletes who are interested in their performance - the amateur athlete and occasional gym attendee often wants to get the most from their work-out and achieve their own personal fitness goals. Products are already appearing that target this emerging market. For example, the Polar heart rate monitor helps people to train at the right intensity depending on their training plan [3]. Polar also offer the service of developing personal training plans based on the data gathered. Adidas have developed the miCoach concept [4] which also measures heart rate data and suggests various plans to help people to train for their specific goals e.g. to lose weight, run a race, de-stress or run faster. Nike and Apple have developed a sports kit which measures speed and distance while walking or running. The data can then be downloaded onto a personalized homepage. Not only does this allow the user to assess their own performance but it has led to a web community where users can compete against and challenge each other in addition to discussing training and related issues. This has led to a virtual running community with close to 2 million members [5]. The advantage of this system is that it avails of apparel and a wearable/portable device that runners are already using. In this way it requires little or no extra effort on the part of the user. This type of strategy is ideal for all wearable sensing i.e. to customize apparel and on-body devices that people routinely use and augment their functionality. In this way, a truly innocuous means of sensing can be achieved that needs no additional appliances. To do this it is necessary to integrate sensors into textiles so that they are comfortable to wear. This is a challenging task considering the rigor that our clothes go through day to day. Sensors must be robust and reliable being able to withstand stretch, pressure and varying temperatures, environmental conditions and ideally the washing machine. In this paper we present a number of prototype systems which use textile based sensors to monitor the wearer. The first uses a glove that has been developed for a clinical study in stroke rehabilitation. The second uses a shirt to monitor breathing patterns. This is used to provide feedback on breathing technique and encourage patients to perform their breathing exercises. The third is a smart insole which is being developed to investigate running technique. Before presenting the individual prototypes an overview of the textile sensor integration strategies is given in the next section.

2. FUNCTIONALISED FABRICS

In order to monitor the body in a natural way there is a need for integrated sensors that are comfortable, wearable and straightforward to use. A complete system that incorporates wearable sensors and body sensor networks within a textile will require a number of functionalities to be added to the textile structure, including conductivity, sensing, actuation, data transmission and computation. Data transmission is essential between components and also wireless connectivity is often desirable. This is possible in a garment through flexible polymer or textile antennas. [6, 7] Conductivity may be added through conductive threads, fibres or coatings. Stainless steel or silver threads are available in different thicknesses, both as 100% metallic fibres, or blended with other textiles such as nylon, giving rise to a variety of yarns with different properties. [8] There are also polymers which have inherently conductive properties which can be coated onto yarns or fabrics. These materials can be used to create sensors with piezo-resistive properties. Coating a stretchable fabric with a conducting polymer converts the fabric into a strain gauge, as the resistance depends on the degree to which it is stretched [9]. Coating a compressible textile such as polyurethane foam in the same manner produces a pressure sensor [10]. In these cases the textile itself becomes the sensor – thereby making use of the structures/materials that are already in place, but improving on them by giving additional functionality. Ideally the sensor should retain its sensitivity over the lifetime of the garment and through numerous wash cycles. Making the fabric itself the sensor can augment garment functionality while still maintaining the normal tactile properties of the garment. Textile based sensors which are compatible with textile manufacturing processes are essential for such technology to become accessible. This may involve screen printing of thermochromic dyes, knitting conductive yarns or weaving of plastic optical fibres [11]. The interconnection of the sensors to a microcontroller or wireless device still remains an issue as it is at this point where a textile sensor must be connected to more conventional electronics. Flexible circuit boards and batteries [12, 13] are a possible solution and a sewable microcontroller (Lilypad Arduino) has recently been developed which allows conductive threads to be stitched to the pins of the microcontroller [14]. Recent studies have reported textile based transistors using organic field effect transistors which would allow fully computational electronic textiles [7, 15]. Smart nanotextiles are being developed to overcome the shortcomings of interconnections with conventional silicon and metal components which are incompatible with the soft textile substrate[16]. By integrating technology at the nanoscale level the tactile and mechanical properties of the textile may be preserved, retaining the wearable flexible necessary characteristics that we expect from our clothing. Through nanoscale manipulation intelligent textiles are given new functionalities including self-cleaning, sensing, actuating and communication [16]. Nanotechnology allows the incorporation of new functionalities at various stages of production – at fiber spinning level, during yarn/fabric formations or at the finishing stage. Nanocoatings are being widely applied at present by the textile industry to improve textile performance by adding antimicrobial and self-cleaning effects. Other applications under development involve the use of conductive materials such as graphite nanofibers and carbon nanotubes to bring conductivity and anti-static behaviour to the textile [17].

The prototypes presented in the following sections make use of piezoresistive textile sensors. The glove and the shirt use a carbon-loaded elastomer which can be coated onto textiles. The insole makes use of a conducting polymer based felt which acts as a pressure sensor. Bekinox stainless steel thread is used to connect the sensors while conventional components and microcontrollers are used for signal conditioning and data acquisition.

3. CASE STUDIES

3.1 Rehabilitation Glove

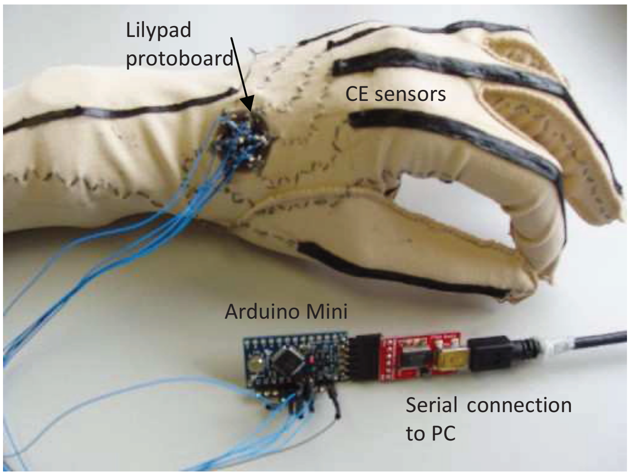


Figure 1 Sensorised glove for stroke rehabilitation

A sensorised glove has been developed for rehabilitation applications to monitor the ability of patients to perform finger extension movements during stroke rehabilitation. An oedema glove, which is typically worn by patients following a stroke to reduce swelling of the hand, has been modified to integrate bend/stretch sensors to measure finger movement, as shown in Figure 1. The glove is made from a lycra and spandex material, which fits to the hand closely with light compression. Movement of the fingers causes the fabric to stretch and regions of the finger and thumb of the glove have been coated with a sensing element , a carbon-loaded elastomer which is commercially available from WACKER Ltd (Elastosil LR 3162 A/B). This approach has been demonstrated in the University of Pisa for measuring body kinematics [18]. This glove has been developed for rehabilitation applications based on the Fugel-Meyer Assessment of motor recovery after stroke. This assesses various motor functions and scores them based on performance where 0 = cannot perform, 1 = performs partially, 2 = performs fully. The advantage of this sensor is that it is comfortable to wear and is integrated into a garment that is already being used by many patients. It is important that the sensor does not constrain or enhance the natural movement of the joints. The sensor's performance has been compared to a commercial bend sensor (Spectra symbol Flex sensor 4.6″) which is a more rigid structure. The two sensors are shown side by side in Figure 2. However, the CE sensor has the advantage of being much more flexible and can be made in any dimension or form while the flex sensor is available at a specific length. Figure 3 shows the response of the sensors positioned on the fingers of a glove during three stages of finger extension. Both sensors achieve a similar response although there is some latency with the CE sensor due to the nature of the material.

The elastomer is a two part compound that is mixed together thoroughly before being cured. The substance was coated onto the fingers and thumb of an oedema glove. The glove was then placed in an oven at 80ºC for 2 hours. Conductive thread was used to connect the sensors to a Lilypad protoboard. This is a circular prototype board (2cm diameter) which allows interconnection by embroidering with thread and also allows wires and components to be soldered. An Arduino Mini microcontroller is used to collect the data using an analog input, sampled at 10Hz. This microcontroller has been chosen for its small size. While it does not have wireless connectivity built in, it is possible to connect a BlueSMiRF modem which works as a (RX/TX) pipe through which a serial data stream can be passed seamlessly at baud rate of 9600bps to a laptop up to 30m away.

Figure 2 (top) CE sensor coated onto stretch knit fabric, (bottom) commercial bend sensor

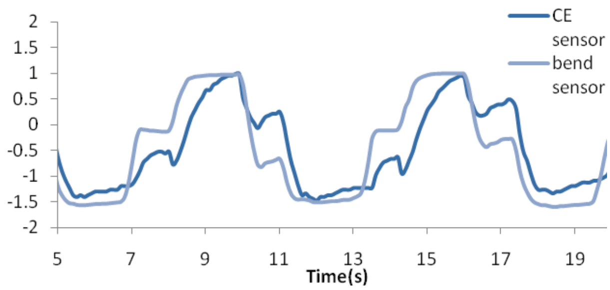


Figure 3 Performance of CE sensor compared to a commercial bend sensor during finger flexion/extension movements

Monitoring joint flexion as demonstrated here with the rehabilitation glove has the potential to automatically assess the patient's performance. A user interface designed for home use on the home PC is currently being developed to determine the position of the user's hand and to present this as visual feedback to the user. The program saves this in an animation format which can then be sent to the therapist who can play back and assess the patient's performance remotely. Figure 4 illustrates the application of this system in a clinical context. The system will be evaluated in forthcoming clinical trials with stroke patients.

Figure 4 System architecture of the stroke rehabilitation process using a sensorised glove

3.2 Breathing Feedback System

Breathing affects virtually every part of the body. It oxygenates the body, revitalizing organs, cells and tissues. Breathing is controlled by the autonomic nervous system and is unique as it is both a voluntary and involuntary process. Good breathing technique can have a profound effect on overall performance as athletes, singers, and yoga practitioners know. In the case of conditions such as asthma or cystic fibrosis it is important for patients to regularly exercise their lungs and improve their breathing technique. Exercising respiratory muscles can increase exercise capacity and reduce the aspects of breathlessness.

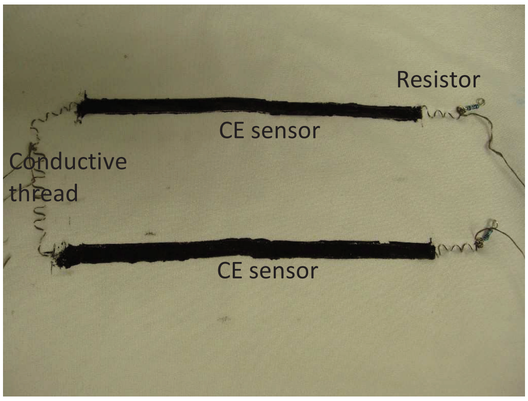


Figure 5 Carbon-elastomer sensors coated on stretch fabric connected by conductive threads. Resistor pins have been bent to allow connection by embroidery of conductive thread.

One exercise that is widely accepted is known as the Active Cycle of Breathing which uses breathing exercises to remove phlegm from the lungs. Clearing secretions from peripheral airways is the most important defence mechanism of the respiratory system. The technique involves taking 4-5 deep breaths with holding periods in-between to allow air to be transported behind obstructed areas in the lungs [19]. It is important that the patient breathes from the diaphragm and not just the thorax for the technique to be effective. For patients with cystic fibrosis this may need to be carried out three times a day. It would be of great use to develop an interactive system which would help children with cystic fibrosis to perform these exercises correctly, and to provide an incentive to perform the tasks regularly. Therefore we have developed a system which uses a wearable sensor to detect breathing patterns, records the signals and provides immediate feedback to the user. As with the stroke rehabilitation example (above), this system may be used at home, while also facilitating remote supervision by a trained therapist who can provide regular assessment updates.

Breathing rates can be measured by detecting the expansion and contraction of the ribcage, and fabric strain or pressure sensors can be used to measure this movement. Carbon-loaded elastomer sensors, as shown in Figure 5, have been used to develop a shirt to measure the expansion of the thoracic/abdominal cavity. The advantage of using these sensors is that they can be applied in any dimension to any desired position on the fabric. In order to maximize the signal quality, the sensors need to be coupled closely with the body movements, and therefore the garment needs to fit the wearer well. For greater flexibility a number or chest straps have also been developed for a universal fit during clinical trials. Data is collected using an Arduino Mini microcontroller as used for the glove discussed in Section 3.1.

The signal is filtered using a low-pass Butterworth filter and the filtered signal is cross-correlated with a reference signal to assess the performance of the user. A Flash application is used to present feedback to the user in a graphical format. The use of a Flash application makes the software easily accessible and means that the system has the potential of being accessed online where the results can be automatically logged to a user's webpage and also accessed by the therapist. The user interface presents an avatar in a picturesque setting as shown in Figure 6.

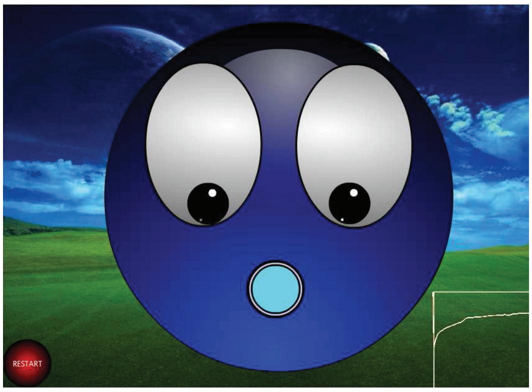


Figure 6 Graphical User interface, Avatar gives instruction and real-time feedback to the user

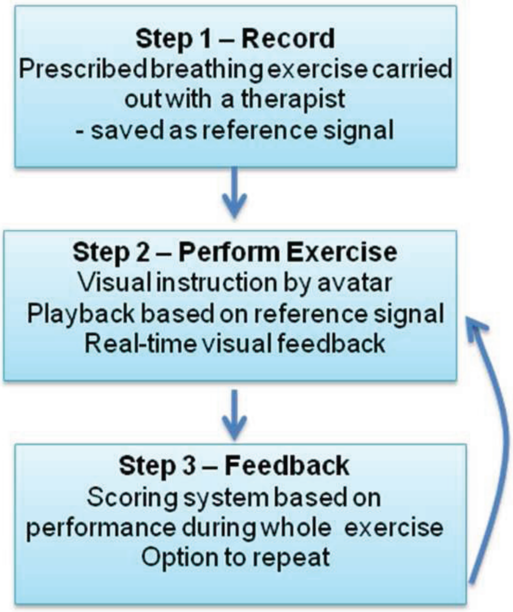


Figure 7 Breathing feedback system operation

The avatar serves to instruct the user while also giving real-time feedback of progress. The mouth of the avatar expands and contracts as the user breathes in and out. An overlay ghost image of the mouth represents the user's breathing. Therefore as they perform the exercise they must try to follow the breathing sequence as instructed by the avatar and therefore synchronise the "ghost" mouth and avatar's mouth. An overall assessment of their performance throughout the exercise is given on completion of the exercise. Rather than giving a numerical score three, four or five stars are given. The idea is that the feedback should be rewarding and encourage the user to keep performing the exercises regularly.

The sequence of the program operation is displayed in Figure 7. The first time that the exercise is performed the signal is recorded as a reference. This is designed to be performed under the supervision of the therapist so that the correct procedure is followed. The avatar is not displayed during this time, but the signal is graphed to show that everything is connected and working correctly. After this, patients may repeatedly perform the exercise in their own time.

To test the effectiveness of the avatar a breathing exercise was performed involving slow deep breathing at a rate of 8 breaths/min for 2 minutes. Breathing patterns were recorded using Sensormedics Vmax as a reference in addition to signals measured at the chest and abdomen using the textile sensors. Figure 8 and Figure 9 show the effect of the breathing exercise under the avatar's instruction. The breathing rate measured by the Sensormedics system changes from a range of 6 to 18 breaths per minute to a constant 7-8 breaths per minute once the exercise is started. Figure 9 shows an increase in amplitude in the abdominal breathing signal during the exercise as the subject is taking deep breaths and using the full lung capacity.

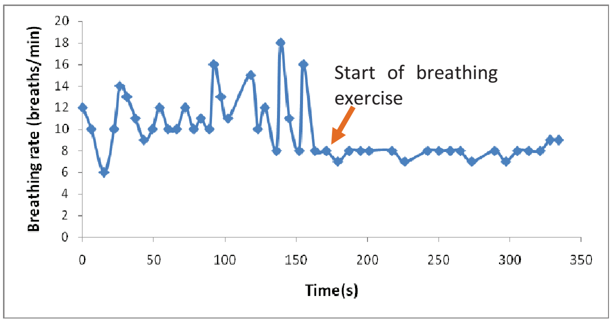
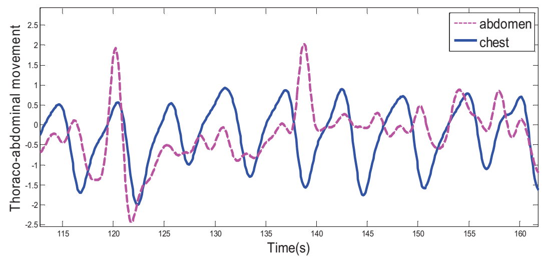


Figure 8 Breathing patterns measured using Sensormedics Vmax before and during breathing exercises using avatar instruction

(a)



(b)

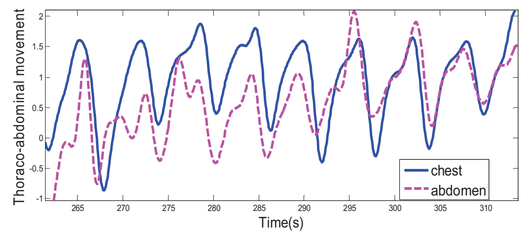


Figure 9 Breathing patterns measured from textile sensors placed at the chest and abdomen a) before breathing exercise b) during breathing exercise

3.3 Smart insole

A smart insole has been developed to detect footfalls during walking or running. The sensor used is a piezoresitive felt produced by Eeonyx (a non-woven textile coated with conducting polymer) whose resistance depends on the applied pressure. .The insole has been made by sandwiching this piezoresistive material at various locations between two layers of neoprene that have been cut into the shape of an insole. Figure 10 shows either side of one of the neoprene layers. Connection to the sensor is made using conductive thread. This is placed in the bottom bobbin case of a sewing machine while non-conductive thread is used as the upper thread. The purple, outer side of the insole has non-conductive thread stitching while the conductive thread is on the inside in contact with the sensor material. Sensors are placed at the heel and toe and at the left and right forefoot. Heel and toe sensors are important for investigating foot contact times while the forefoot sensors at each side help to determine the roll of the foot. Often runners are subject to injuries based on supination or pronation of the

Figure 10 Neoprene insole with conductive stitching for the placement of four pressure sensors

foot when the foot rolls to the outside or the inside [20]. The insole may be used for various applications to assess athletic performance. It may be used to look at how foot contact time changes throughout a track race or during long distance events. It can be used to build an archive of running history and follow the effects of different training regimes on running technique. It allows every footfall to be captured and recorded to keep track of training achieved and to help devise future training strategies. It may be used to prevent or detect injuries if an unusual pattern emerges.

Apart from sports applications, this type of wearable device may prove useful as a tool for gait analysis in rehabilitation, such as monitoring of the progression of Parkinson's disease, or activity monitoring in the elderly. The insole is comfortable to wear and can be inserted into any type of shoe. The placement and number of sensors can be adapted to the particular needs of the application.

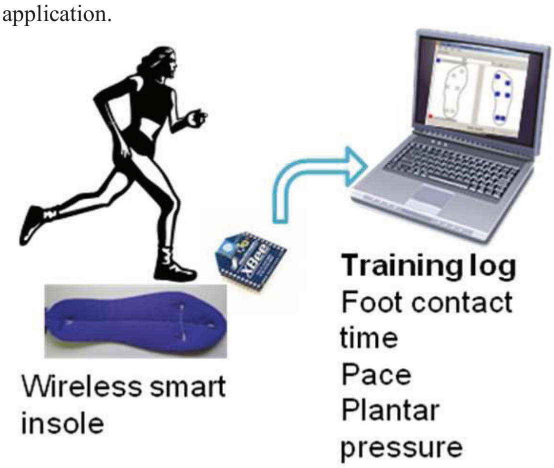


Figure 11 Smart insole to monitor running technique

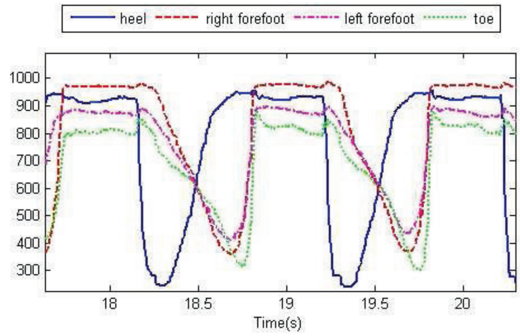


Figure 12 Response of the smart insole during walking

Data is acquired using an Xbee module (1mW chip antenna). This offers improved data acquisition performance over the Arduino platform used in the previous two examples. The Xbee is held in position by a neoprene ankle strap attached to the insole. The Xbee has 6 channels which can be conditioned as analog or digital inputs. Sampling rates up to 1kHz across all channels is possible. For the insole with four sensors each channel has been set to sample at a rate of 200Hz. An Xbee regulated USB module (Sparkfun Electronics) is used to connect a base-station receiver Xbee to a laptop which graphs the data in real-time and stores the data for analysis.

The response of the insole is shown in Figure 12 during walking. From this it is clear about the timing as each part of the foot strikes the ground. Time between steps can be estimated by measuring the time between heel strikes. Foot contact time can be estimated by measuring the time between heel strike and toe off.

4. DISCUSSION/CONCLUSIONS

Fully integrated wearable sensing technologies should be based on materials that are soft, flexible and washable, to meet the requirements of normal clothing manufacturing. Data transmission must be wireless to allow free movement of the wearer, and ideally make use of technology that is already being carried or worn by the wearer if it is not automatically built into the garment. Clearly sensor integration and signal processing are significant challenges, but one of the biggest questions is what to do with the information that is gathered. It needs to be presented to the user in a suitable feedback approach e.g. visual, tactile or auditory and the timing and frequency of this feedback is critical. This is largely application dependent and dynamic in nature, and context awareness is therefore a crucial feature in generating user feedback. The feedback should be intuitive, without overloading the wearer with too much unnecessary information. Put simply, the right information needs to be given to the wearer, at the right time. Wearable sensors can help to achieve a better understanding of personal physiology and, combined with contextual information, they can help to create a more complete archive of a person's lifestyle.

Three different systems involving textile based wearable sensors to detect body movements have been presented in this paper. Using such wearable devices, it is possible to acquire information about the person in a natural setting without intruding on their daily life. This is made possible by augmenting the functionality of the type of clothing that is commonly in use for each application. In this way it is possible to build an archive of personalized information about the wearer's activities, for example, to track progress in prescribed rehabilitation exercises or fitness training plans. By extracting digital information from the threads of our clothing it is possible to build a digital archive relating to our own body's physiology. The coming decade will clearly see a rapid expansion in the numbers of people using wearable sensors for many applications in Health and Exercise. The availability of data in archived databases opens the way to comparing activities of people on a global scale, as has been seen with the Nike plus system where thousands of people are collectively running millions of miles. It also allows for personalization of systems in addition to giving us a better understanding of our personal lifestyle and health, and how we compare to others. While this undoubtedly have many benefits in highlighting global trends, the ownership of this type of data still needs to be regulated and the privacy of individuals must be respected. Novel sensing technologies enable us to harvest large volumes of sensed information which must be used carefully and productively to benefit the whole population.

ACKNOWLEDGMENTS

We gratefully acknowledge the financial support of Science Foundation Ireland (07/CE/I1147) for this research. We are grateful to the staff at the William Stokes Unit in the Adelaide and Meath Hospital, Tallaght, Dublin 24, Ireland for their help.

REFERENCES

1. Paradiso, R., G. Loriga, and N. Taccini (2005) A wearable health care system based on knitted integrated sensors. IEEE Trans Inf Technol Biomed. 9(3):337-344.
2. Heilman, K.J. and S.W. Porges (2007) Accuracy of the LifeShirt® (Vivometrics) in the detection of cardiac rhythms. Biological Psychology. 75(3):300-305.
3. Polar http://www.polar.fi/en (accessed Jan 2010).
4. Adidas miCoach http://www.adidas.com/uk/micoach/#Start/sdf/mdf (accessed Jan 2010).
5. Nike Running http://nikerunning.nike.com/nikeos/p/nikeplus/en_US/ (accessed Jan 2010).
6. Locher, I., M. Klemm, K. T, and G. Tröster (2006) Design and Characterization of Purely Textile Patch Antennas. IEEE Transactions on Advanced Packaging. 29(4):777-788.
7. Declercq, F., H. Rogier, and C. Hertleer Permittivity and loss tangent characterization for garment antennas based on a new matrix-pencil two-line method IEEE Transactions on Antennas and Propagation. 56(8):2548-2554.
8. Post, E.R., M. Orth, P.R. Russo, and N. Gershenfield (2000) E-broidery: Design and fabrication of textile-based computing. IBM Systems Journal. 39(3&4):840-859.
9. Brady, S., S. Coyle, Y. Wu, G. Wallace, and D. Diamond.(2006) *Wireless-based monitoring of body movements using wearable sensors.* In: Proc. *MRS Spring Meeting, Symposium S*. 2006. San Francisco.
10. Brady, S., D. Diamond, and K.T. Lau (2005) Inherently conducting polymer modified polyurethane smart foam for pressure sensing. Sensors and Actuators A. 119:398-404.
11. Braddock Clarke, S. and M. O'Mahony, *Techno Textiles 2: Revolutionary Fabrics for Fashion and Design* 2006: Thames and Hudson.
12. Nishide, H. and K. Oyaizu (2008) Materials science. Toward flexible batteries. Science. 8(319):737-8.
13. Siegel, A.C., S.T. Phillips, M.D. Dickey, N.L. Zhigang Suo, and G.M. Whitesides (2010) Foldable Printed Circuit Boards on Paper Substrates. Advanced Functional Materials. 20(1):28-35.
14. Buechley, L. and M. Eisenberg (2008) The LilyPad Arduino: Toward Wearable Engineering for Everyone. IEEE Pervasive. 7(2):12-15.
15. Maccioni, M., E. Orgiu, P. Cosseddu, S. Locci, and A. Bonfiglio (2006) Towards the textile transistor: Assembly and characterization of an organic field effect transistor with a cylindrical geometry Export Applied Physics Letters. 89(14):1-3.
16. Coyle, S., Y. Wu, K. Lau, D. De Rossi, G. Wallace, and D. D (2007) Smart nanotextiles : A review of materials and applications. MRS Bulletin. 32(5):434-442
17. Qian, L. and J.P. Hinestroza (2004) Application of Nanotechnology for high performance tectiles. Journal of Textile and Apparel, Technology and Management. 4(1):1-7.
18. Carpi, F. and D. De Rossi (2005) Electroactive Polymer-Based Devices for e-Textiles in Biomedicine. IEEE Transactions on Information Technology in Biomedicine. 9(3):295-318.
19. Hristara-Papadopoulou, A. and J. Tsanakas (2007) Results of active cycle of breathing techniques and conventional physiotherapy in mucociliary clearance in children with cystic fibrosis. Hippokratia. 11(4):202-204.
20. Goldberg, D.A. and D.L. Whitesel (1983) Heel counter stabilization of the running shoe. The Journal of orthopaedic and sports physical therapy. 5(2):82-83.

FoodLog: the Easiest Way to Capture and Archive What We Eat

Kiyoharu Aizawa, Gamhewage C. de Silva, Keigo Kitamura, Yuto Maruyama
The University of Tokyo
{aizawa, chamds, kitamura, maruyama}@hal.t.u-tokyo.ac.jp

ABSTRACT
We present the current status of our web-based social multimedia application that can be used as a dietary management support system by ordinary Internet users. It analyzes image archives that belong to the user to identify images of meals. Further image analysis determines the nutritional composition of these meals and stores the data to form a Foodlog. The user can view the data in different formats, and also edit the data to correct any mistakes that occured during image analysis. This application is currently open to the public and has approximately 280 users, who have uploaded more than 5500 images of meals. We present the current status of this application, and discuss our future plans to extend it to allow interaction between users for more effective dietary management.

Categories and Subject Descriptors
H.4 [Information Systems Applications]: Miscellaneous; H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval

Keywords
Foodlog, life-log, dietary management, multimedia

1. INTRODUCTION
With obesity becoming a growing health concern in many parts of the world, dietary control has received considerable attention in the field of healthcare. However, most dietary control programs require manually logging detailed information regarding all meals. This is a tedious task for an ordinary person, and can be a deterrant to taking part in such programs.

With the widespread use of digital cameras and camera phones, a person now has easy access to a camera during most of his daily activities. Therefore, one can use photos of meals as a record of one's dietary information. Images contain much richer information than a textual or tabular description of meals. A recent study shows that merely taking a photograph of a meal before eating can encourage weight loss [5]. Therefore, images are highly prospective candidates for fast and easy recording of dietary information. However, they also present an additional problem; images need more time to analyze than a textual meal summary.

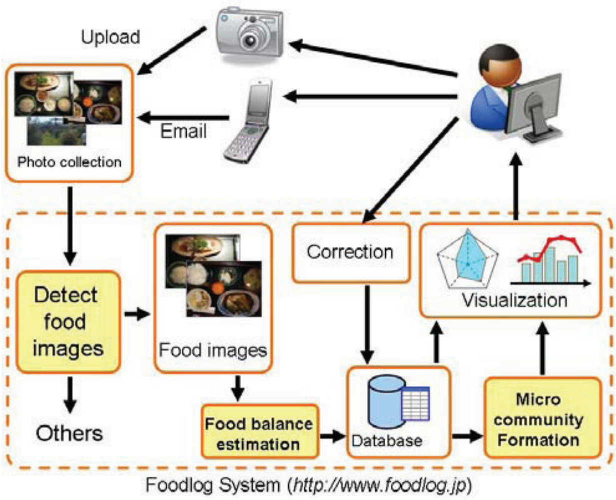


Figure 1: System overview.

While ease of recording dietary information is important, it is not always sufficient for effective dietary management. A participant of a dietary program regularly meets a consultant with the collected data, to get the data analyzed and receive recommendations based on them. If this step can be made more regular, but at the same time less costly and time-consuming, the results can be greatly improved. Another important aspect is finding "peers" who can join a participant on dietary control, for interactions such as encouragement and informal advise. It is common for people to look for a partner or groups to participate in activities that need a long time effort, so that they can keep their motivation.

In the light of the above observations, we propose a web-based multimedia application that can assist an ordinary person to easily record and manage his/her dietary activities with minimal effort. Figure 1 outlines the functionality of this application. Instead of creating a detailed records

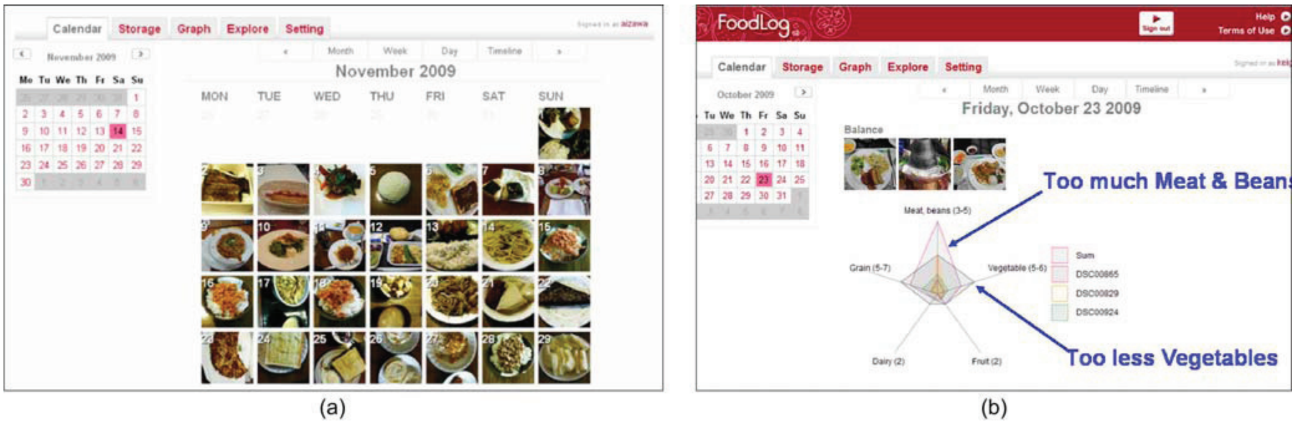


Figure 2: Summaries on Foodlog: (a) a monthly meal summary with a calendar-based visualization; (b) the photos and dietary balance for a single day.

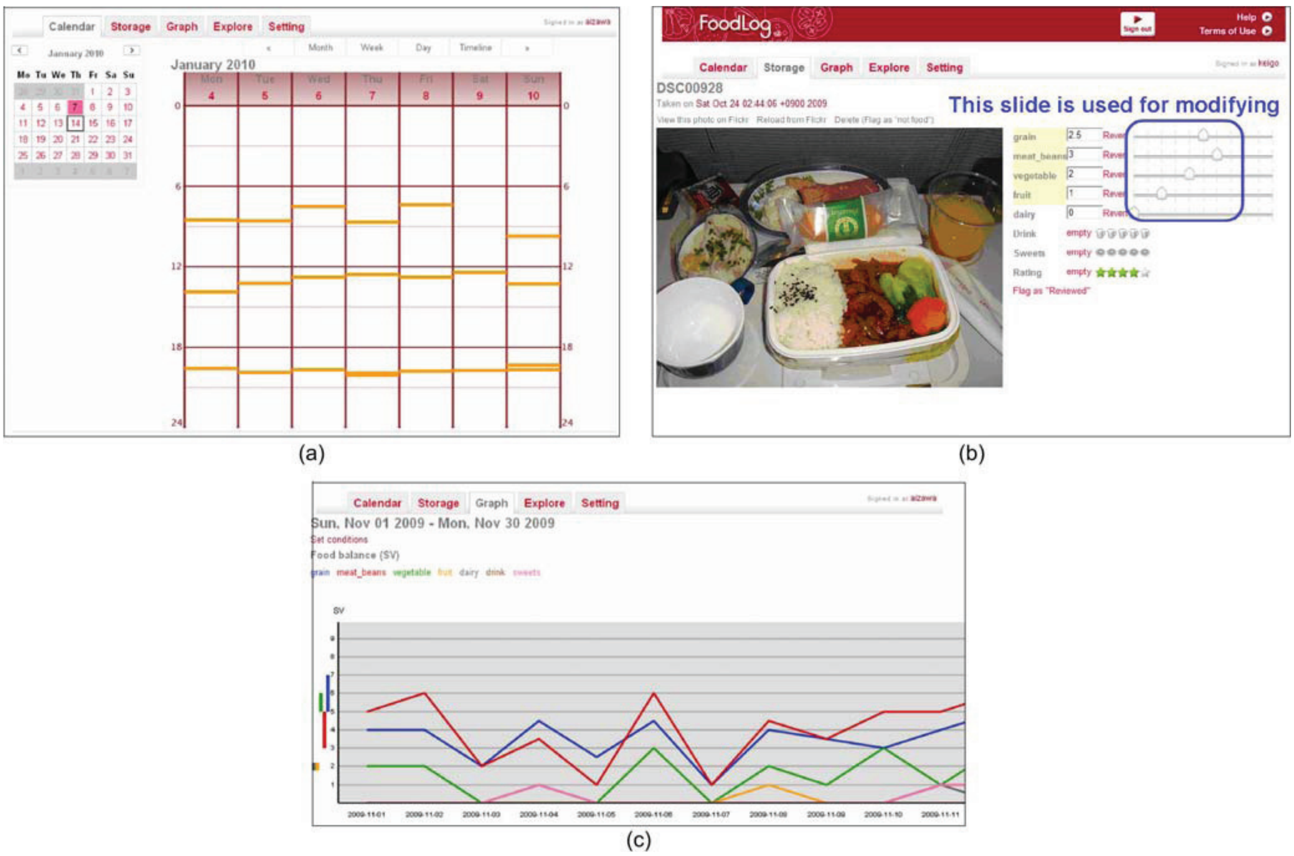


Figure 3: Additional screenshots: (a) weekly timeline of meals; (b) the interface for managing attributes of a given meal and correcting dietary balance; (c) graph showing variation of dietary balance with time.

of each meal, the user simply takes a photo using a digital camera or a camera phone. The application uses image analysis to detect images of meals in his personal multimedia archive, eliminating the need for the user to select and upload them. Further image analysis estimates the nutritional composition of the meals, and records the results in a database. The user can access the images and the results on the web, in various formats. He/she can also revise the information where image analysis has been erronerous, and add extra information where relevant.

We also propose to form microcommunities by analysing data from different users and finding similarities, so that they can interact with, support and be supported by other users. The following sections of this paper reports the current status of this application, and outlines future research directions.

2. SYSTEM DESCRIPTION

The user records his dietary information by taking photos of his/her daily meals with a digital camera and/or a camera phone. These photos are uploaded to the user's personal Flickr account, together with other photos. For photos taken with a camera phone, it is easy to upload images by sending them to Flickr via email. The proposed system crawls this account regularly to look for new images, and then analyzes them to identify images of meals. An algorithm based on Support Vector Machine Learning of both local and global image features [2] is used for this purpose.

The dietary balance of the meals is calculated according to the "Food Balance Guide" of the Ministry of Agriculture, Forestry and Fisheries of Japan [1] , which is similar to "My Pyramid Specification" by the U.S. Department of Agruculture [4]. It categorizes food into five groups: grains, vegetable, meat/beans, milk, and fruit. The quantity of food in each group is defined by the original unit, serving (SV). While not as detailed as the number of calories, this still provides a reasonable description of dietary balance of a meal. An algorithm based on color histograms and DCT coefficients and SIFT features [2][3] is used to estimate the dietary composition of each food image. The results are stored together with the date and time the photo was taken. Location data, if available with the image, are also included. The collection of thus recorded data form a log of the user's meals.

The user can browse the results by logging on to the application using a web browser. At the current state, the interfaces are designed for a computer. The user can view photos organized in daily, weekly and monthly formats. Figure 2a shows a screen capture of a monthly summary of food images. A summary of food intake for a single day is shown in Figure 2b, with a chart showing the daily intake of each nutritional category. Figure 3a shows shows the timeline of meals during a given week. This visualization enables a user to quickly identify patterns in his/her meal times. In addition to these visualizations with fixed time intervals, a user can specify a range of dates and obtain a line graph showing the diatary balance for that range (Figure 3c).

The system also allows a user to correct any errors in the results of image analysis. The user can drag and drop images between the two categories "meals" and "non-meals". A graphical interface with sliders is used for making corrections to the dietary balance (Figure 3b).

At the current state of the system, there is little interaction among different users. A user can assign a rating to a meal that he took, while browsing his foodlog. The system provides an option for users to "explore" rated images by other users. This indirectly helps them to increase their awareness about meals and there compositions.

3. INTERMEDIATE RESULTS

At the time of writing this paper, the system is open to the public under the name Foodlog, with the functionality described in Section 2. Currently, Foodlog has 280 subscribers and approximately 5500 images of meals.

The content uploaded by users demonstrate the effectiveness of providing users an easy method to record their dietary activities. For instance, one of the authors was able to record more than a year of his dining history with very little effort. Another user, who is conducting research on this system, manually recorded his meal information for 35 days and compared it with the content uploaded to Foodlog. It was found that 91% of all manually recorded information was contained in images uploaded to Foodlog.

4. TOWARDS A COMPLETE DIETARY SUPPORT SYSTEM

The algorithms for image analysis can be improved to achieve more accurate food image recognition and balance estimation. We are now conducting initial studies on automated calorie estimation. User interaction for mobile devices should be desgned so that the results too can be easily accessed.

At the current state of Foodlog, it is a multi-user system for individual management of dietary information. We are working on increasing the interaction between subscribes users so that "microcommunicites" can be formed to support users.

5. REFERENCES

[1] Food balance guide. ministry of agriculture, forestry and fisheries of japan. http://www.maff.go.jp/j/balance-guide/.

[2] K. Kitamura, T. Yamasaki, and K. Aizawa. Food log by analyzing food images. Proceeding of the 16th ACM international conference on Multimedia, pages 999–1000, 2008.

[3] K. Kitamura, T. Yamasaki, and K. Aizawa. Foodlog: Capture, analysis and retrieval of personal food images via web. ACM Multimedia Workshop on Multimedia for Cooking and Eating Activities, pages 23–30, 2009.

[4] MyPyramid. United states department of agriculture. http://www.mypyramid.gov/.

[5] L. Zepeda and D. Deal. Think before you eat: Photographic food diaries as intervention tools to change dietary decision making and attitudes. International Journal of Consumer Studies, 32(6):692–698, November 2008.

Evaluating Interactive Access to Personal Visual Archives

Kan Ren
I-Lab
University of Surrey
Guildford, UK
k.ren@surrey.ac.uk

Risto Sarvas
Helsinki Institute for
Information Technology
PO Box 9800, 02015 TKK,
Finland
risto.sarvas@hiit.fi

Janko Ćalić
I-Lab
University of Surrey
Guildford, UK
j.calic@surrey.ac.uk

ABSTRACT

Triggered by a surge in amount of personal multimedia content, both stored locally or shared online, the systems for organisation, search and browsing of user-generated visual media have become increasingly important. This work presents a number of user studies into effectiveness of interactive search and browsing of personal photos and videos. The study is conducted using a novel intuitive interactive interface for browsing of large-scale image and video collections. The system clusters the dataset by exploiting a highly efficient unsupervised graph-based clustering algorithm and visualises the underlying dataset structure by the size and spatial relations of displayed images. The extensive experimental results of task-based user studies into search and selection of personal photo collections demonstrated high level of user satisfaction and good usability of the system when compared to the existing systems for access to personal multimedia archives. Contrary to the task-based experiments, the experience of the studies of spontaneous usage of personal archives introduces overwhelming challenges in quantifying experimental results. Nevertheless, the subjective qualitative evaluation consistently revealed positive user experience of cross-user interaction in photo sharing when using the proposed interface.

Categories and Subject Descriptors

H.3.1 [**INFORMATION STORAGE AND RETRIEVAL**]: Content Analysis and Indexing—*Abstracting methods*; H.3.7 [**INFORMATION STORAGE AND RETRIEVAL**]: Digital Libraries—*User issues*; H.5.2 [**INFORMATION INTERFACES AND PRESENTATION**]: User Interfaces—*Graphical user interfaces, Screen design, User-centred design*

General Terms

Algorithms, Performance, Design, Experimentation, Human Factors

Keywords

User study, personal photo collections, image browsing, HCI, image clustering

1. INTRODUCTION

The dramatic proliferation of personal digital content, generated by users themselves as well as the omnipresent capture devices that surround us, has transformed the way content is maintained, managed and exploited. Driven by the continually changing environment and the need for effective management of large-scale multimedia datasets, there is a strong demand for efficient and flexible way of interaction with the digital content. Personal media devices such as digital camera or video recorders have become a commonplace. Users can easily take hundreds of photos and video clips on a daily bases, but only a few generate meaningful annotations to facilitate automated structuring and management of those large personal media archives. This implies that the user's local storage is filled with photos and video clips in an unordered manner. The problem of structuring, browsing and retrieving content from such collections is becoming a major challenge of multimedia management systems.

The two major existing approaches that tackle this problem either require incremental and interactive manual annotation, or generate annotation automatically using content-based media analysis, computer vision and machine learning. However, due to users reluctance to manually label large quantities of data, of the problem of semantic gap between the low level features and high-level semantic understanding of the media, the automated annotation has not deliver satisfying results.

The system presented in this paper makes a shift towards more user-centred design of interactive image and video search and browsing by augmenting user's interaction with content rather than learning the way users create related semantics. This shift enables not only efficient retrieval of the desired content, but offers more intuitive access to vast visual data and often gives unexpected perspective of the explored dataset. Finally, this approach facilitates more intuitive and effortless browsing, enabling exploitation of the system by a wider user base.

The paper presents four user studies of the search and browsing interface, incorporating experiences of personal photo and video browsing and retrieval, as well as cross-user interaction with large-scale shared personal archives. The results show very positive user experience when using the proposed interface, especially while browsing or selecting a subset of content from the archive. In addition, the user study brought difficulties in quantifying evaluation results of the spontaneous usage of the system, while the results of task-based experiments were robust and consistent.

The work related to this research area is presented in Section 2, while the Section 3 brings the methodology used in designing and implementing the browsing interface. In order to evaluate the presented system Section 4 describes the experimental setup, while Section 5 discusses the achieved results and outline the future plans.

2. RELATED WORK

There has been a lot of effort put in the scientific research as well as commercial development of user-friendly image and video browsing applications. Most of the browsing applications are based on the time domain clustering of the personal photo collections, having the temporal metadata readily available from the digital cameras. The applications simply cluster the images based on the time when the photo or video clip was generated [1] [2] [3]. But the disadvantage of this approach is that the user needs to type manually an event name for a group of photos, which can be inexact and unreasonable given the fact that events can span more groups and vice versa.

Triggered by the proliferation of global positioning system (GPS) technology, some of the new applications are using the image similarity based on the location where the operator took the photos [4] [5]. Being an emerging technology in this context, GPS modules are still rarely built into the camera, so users often need to assign the location information manually.

Recently, some commercial applications introduced semi-automated annotation of images by using the face recognition technology [5]. The application first detects face region in the photo and then attempts to identify and tag the image by using face similarity algorithm. However, this approach is unfeasible for many photos, such as landscape photos, animals, etc., since people are not always the major subjects in a captured scene.

There have been a number of approaches to develop visualisation that would augment the usability of interfaces to large image collections. In [6], Huynh et. al. introduced a method that trades off screen space for better presentation of temporal order in photos. In addition, some systems utilised methodologies to analyse the underlying data structures to present image collections [7] in a more accessible way.

However, the interaction with large visual collections has not been addressed in an intuitive way. Derived from its definition in [8], intuition implies correlation between system inference and the user's expectations. By following this definition, we developed an intuitive interactive interface (dubbed *FreeEye*) for browsing of large image and video collections, based on the efficient image clustering method and interactive hierarchical interface.

In order to facilitate interactive browsing of video content by the means of the *FreeEye* interface, the proposed system efficiently extracts a set of representative key-frames from the sequences present in the repository by unsupervised clustering methodology. There is a number of similar approaches that utilise unsupervised clustering in the process of key-frame extraction. An efficient clustering method has been utilised in [9] [10], where K-means algorithm is used to classify data into a fixed number of groups, starting from a random initial partitioning. In [11], an unsupervised clustering based approach was introduced to select key-frames within predetermined shot boundaries.

Similarity comparison using a shot histogram analysis and subsequent clustering is carried out within each shot to automatically select the most representative key-frames. A comprehensive overview of work focusing on evaluation and user studies of video browsing and summarisation methods was presented in [12]. Focusing on the frame saliency and importance in the video summarisation context, a number of graph-based methods have been proposed [13] [14] [15]. However, the efficiency of these approaches heavily depends upon the size of the dataset, due to a high complexity of the spectral analysis exploited in their graph representation. Nevertheless, there have been proposals to analyse visual similarity in the graph-based context with almost linear complexity to the number of nodes in the graph representation. Developed for efficient image segmentation, the algorithm presented in [16] introduces a graph predicate that keeps the notion of global features while making fast decisions locally.

In order to study implication of the proposed browsing interface to the critical practice of photo sharing [17] and collaborative experience of storytelling [18], this paper brings a detailed study of cross-user experience of the interaction with large-scale personal archives.

3. BROWSING INTERFACE

In order to interactively browse large image and video collections, the browsing interface follows the idea of ranked image representation, where more relevant images should be more apparent and thus displayed bigger. This is supported by a hierarchical layout of images on the screen. When user selects an image from the displayed dataset by clicking, the image is relocated to the centre, while the remaining data is retrieved from the repository and arranged on the screen. By doing this, the user practically moves the centre of perspective from which the collection is explored.

The image browsing system comprises two main modules: image clustering engine and the interface generation, as depicted in the Figure 1. The image rank in a generated display is proportional to the similarity measure between user-selected central image and other images from the dataset. The choice of the similarity metric is completely independent of the proposed clustering engine and interactive interface, enabling generic applications of this system. In this paper we conducted user-tests exploiting two descriptors: an three dimensional RGB colour histogram and the timecode field from the photo's *Exif* data. The video sequences are abstracted by a set of key-frames, efficiently extracted by using our previous work on video summarisation [15].

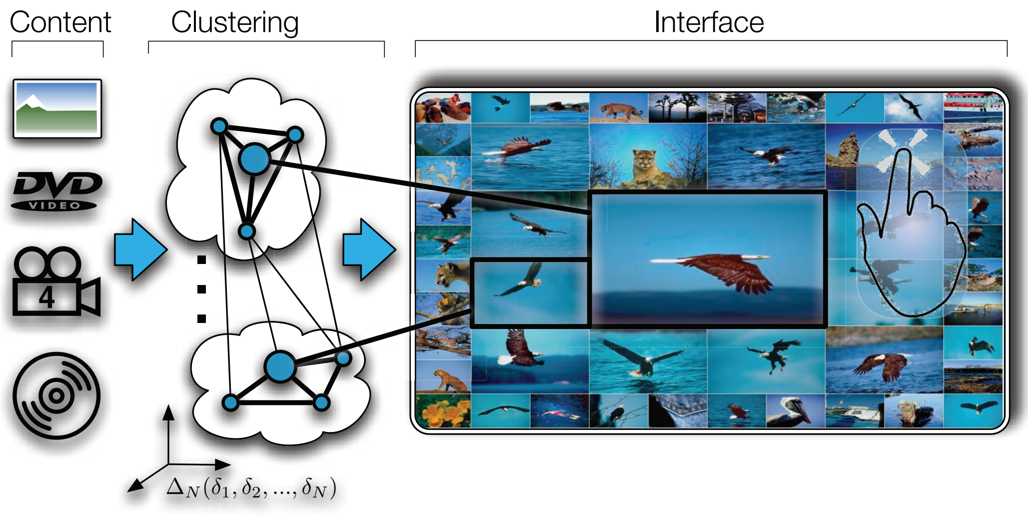


Figure 1: Building blocks of the image browsing interface

3.1 Image Clustering

To achieve system scalability and algorithm complexity nearly linear to the number of key-frames, a graph based clustering algorithm is utilised[16]. Its ability to preserve detail in low-variability clusters while ignoring detail in high-variability regions maintains notion of global features of the dataset in the process of making greedy decisions locally. The graph $G = (V, E)$, is formed so that each image corresponds to a node $v_i \in V$, and the images are connected by undirected edges $(v_i, v_j) \in E$. Weights of each edge $w(v_i, v_j)$ measure the dissimilarity between the two corresponding images.

The graph node grouping is defined by a graph predicate $D(c_1, c_2) : Ext(c_1, c_2) > mInt(c_1, c_2)$, which evaluates if the two clusters c_1 and c_2 should stay disconnected by comparing inter and intra cluster differences as depicted in Figure 2. The internal difference of a cluster c is defined as the largest weight in the minimum spanning tree $MST(c, E)$ of the cluster c:

$$Int(c) = \max_{e \in MST(c,E)} w(e) \quad (1)$$

The joint internal difference measure $mInt(c_1, c_2)$ is therefore given as:

$$mInt = min(Int(c_1) + \tau(c_1), Int(c_2) + \tau(c_2)) \quad (2)$$

The external difference between two clusters $Ext(c_1, c_2)$ is the minimum distance between the two nodes that are members of different clusters:

$$Ext(c_1, c_2) = \min_{v_i \in c_1, v_j \in c_2} w(v_i, v_j) \quad (3)$$

The threshold function $\tau(c) = k/|c|$, where k is some constant parameter and $|c|$ denotes the size of c, controls the degree to which the difference between the two components must be greater than their internal differences. The intra component difference is defined as the minimal weight edge connecting the two components. The technique adaptively adjusts the merging criterion based on the degree of variability in neighbouring regions of the dataset. The node grouping is iteratively repeated until there is no more component merging.

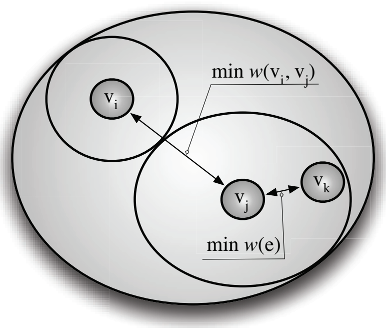


Figure 2: Graph predicate utilised in the clustering module.

3.2 Interface Design

Following the idea of ranked image representation, our aim is to generate an intuitive and easily-readable interface by conveying the significance of an image from the dataset by its displayed size. Any function that evaluates the significance is highly dependent upon the application. In our case, the objective is to facilitate intuitive interaction with large-scale personal archive. Therefore, the interface should clearly present visual content that is dominant throughout the analysed dataset, as well as to present some unanticipated content.

More generally speaking, the task of comprehending large-scale collections solely by the means of interaction and browsing, the user interface needs to achieve a balance between the process that duly favours dominant information and the discovery of the content that is poorly, if at all, represented on the screen. Keeping this balance is important especially in case of visual cognition, where introduction of unantici-

Figure 3: Example of the image browsing interface *FreeEye*

pated visual stimuli can dramatically change the conveyed meaning of represented content. In a series of experiments conducted to indicate the usefulness and effectiveness of film editing [19], Russian filmmaker Lev Kuleshov (circa 1918) demonstrated that juxtaposing an identical shot with different appendixes induces completely different meaning of the shot in audiences. In other words, the conveyed meaning is created by relation and variance between representing elements of visual content. This idea of emphasising difference, complexity, and non-self-identity rather than favouring commonality and simplicity and seeking unifying principles is well established in linguistics and philosophy of meaning through theory of *deconstruction*, forged by French philosopher Jacques Derrida in the 1960s [20].

Following this paradigm, the interactive interface is generated following two main objectives: i) to visually convey data structure extracted in the image clustering stage ii) to achieve intuitive interaction by balancing content representation and knowledge discovery. The interface design follows support of the hierarchical groupings generated by the clustering engine. The centre image is maximised and displayed at 100% of its size. If the user clicks on an image, the image will move to the centre of the refreshed screen, and the remaining display layout will reform in order to represent images in the vicinity of the central image. The immediate neighbourhood is represented with 12 most similar images from the same cluster encircling the central image. These images are displayed at 50% their original size. The next layer encircling the central cluster contains 36 images displayed at 25% size, separated into two parts: four edges and four corners. The 32 images located at the four edges are representing the centres of clusters closest to the central image. To support knowledge discovery and help users locating other areas of interest, four random pictures from the set of unrepresented images are located at four corners of the screen. Every time the user clicks, the system iteratively rearranges all images as described above. An example of the interface is given in Figure 3.

4. USER STUDIES

The proposed interactive search and browsing interface has been evaluated in three different application scenarios. The first scenario comprised an image search task in minimal time, focusing on the overall intuitiveness and efficiency of the system. The second user study comprised three subtasks of selecting a set of personal photos depicting an event, a holiday and the whole year. Finally, in order to investigate usability of the system in a collaborative and sharing context, a cross-user photo selection task was set and evaluated.

4.1 Interactive image search

The image repository used is a selection of cca. 3000 colour images from the Corel image database. In order to test the effectiveness of the search and browsing tool, the database subset includes multiple semantic concepts such as the wild animals (leopard, eagle, fox, etc.), nature scenery (forest, ocean, etc.), historical buildings (western temples, Asian buildings, etc.), portrait, plants (flower, garden, etc), etc.

The subjective tests were conducted by inviting 26 people to join the challenge *Find me a postcard* [21]. The challenge comprised finding 5 images from a set of 3000 only by means of interactive interface described above. Of 26 people involved, 18 persons were male, and 17 had the advanced computer knowledge. All users were using the tool for the first time and the only requirement was to have a basic knowledge of manipulation with a mouse. The gender, racial and cultural diversity of the subjects was balanced.

The task was to find the 5 fixed images in the predefined order. The content of the five images was varied, as presented in the Figure 4. We recorded the full browsing system state for every user step, which included indexes of all images on the screen, their positions, user selection and timing. This has enabled us to fully reproduce the browsing process for each user and analyse achieved results.

Figure 4: Images used as queries in the interactive image search task

The basic statistics of the experimental results shows that the average time for a user to finish the whole experiment is 8 minutes and 20 seconds in 50 mouse clicks. This gives an average of around 100 seconds time and 10 mouse clicks needed for a user to find an image from the database of 3000 images. Assuming that in the case of thumbnail presentation users need to inspect all images from the data set, the average number of images inspected by using the FreeEye tool is 6 times smaller.

In order to evaluate the interface intuitiveness, the user history records were studied. The Figure 5 shows exemplar browsing paths for every task $i = 1, 2, ..., 5$, given as the distance $\delta_i(p)$ between the target image and the selected image at the progress stage p on browsing path from the start of the task until the targeted image was found.

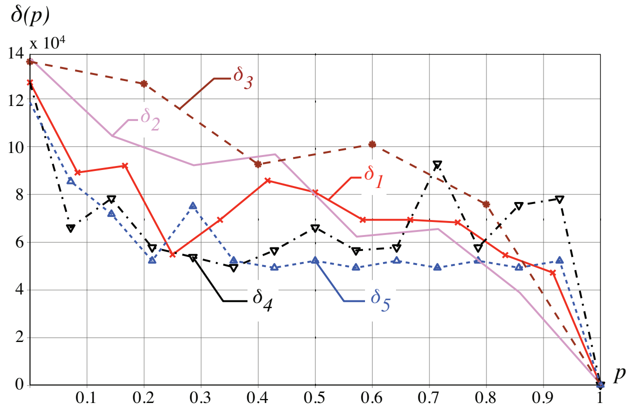


Figure 5: Convergence of browsing paths towards the target images

From all 5 browsing paths, it is observable that after only a few clicks, the distance $\delta_i(p)$ between the current and target image dramatically fall towards zero. This means that the users were rapidly converging towards the goal of the task just after a couple of clicks, implying system's intuitive character. This trend is obvious in the 2^{nd}, 3^{rd} and 5^{th} task, while for the initial task and the 4^{th} task it was more difficult.

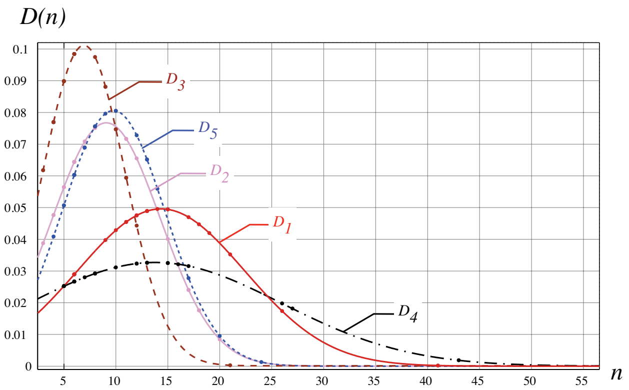


Figure 6: Distribution of clicks for each task in the interactive image search challenge

Since the timing and user clicks directly depend upon the difficulty of the task, we studied the distribution $D(n)$ of the number of user clicks n required to find the target image in the database. From the approximated distribution depicted in the Figure 6, it can be observed that the distributions became increasingly skewed in a positive sense (right-skewed) as the users progress through the tasks. This represents that more users require less iterations to find the desired image as they use the interface. This characteristic demonstrates that without any assistance, users intuitively learn how to efficiently use the interface, regardless of the task difficulty. The same conclusions were made while studying the distribution of time required to find the desired image for each user.

Figure 7: Spatial distribution of user clicks in the interactive image search task

In addition to the click and time statistics, we have studied the spatial distribution of positions of images selected by users. As depicted in the Figure 7, where the region brightness represents frequency of its selection, images in the second level (neighbouring frames of the central image) are selected more often than images in the third level. However, some of the random images in the four corners were occasionally selected, mainly to move away from the currently displayed set of images and test where they would take the user in his search attempt. Furthermore, the top area of the second level was slightly more popular than bottom area, while the right side was a more popular than left side.

4.2 Photo selection task

In order to evaluate the proposed system in a photo selection scenario, we conducted five user trials [22]. The recruited participants were 3 women and 2 men aged 24-32, and all but one had a computer science background. For each trial the participant brought a set of their own digital photos. The number of photos brought by each participant ranged from 1385 to 1664. For each participant there were three separate tasks. The first task was to select photographs from a short-time event (1-2 days) to be sent by email to someone. The second task was to select photographs from a long time event (more than two days) to be uploaded to a web page or shown to someone. The third task was to select photographs for a book representing events and happenings in the past 6-12 months. For each task the participants were asked to think about specific people they would show the photographs to. The selected photographs were not actually sent or shown to anyone outside the trials. After each task the participants were asked a set of questions about the tool, the event, and photographs. The participants were also asked to give a score from 1-5 on how well the tools represented the events, how well the tool helped them to find photographs, and how the tool compared to their regular

	selected photos	time spent	clicks	sec/click	sec/photo
Task 1	10.4	1:52	16.8	6.65	10.7
Task 2	15.6	5:36	49.2	6.82	21.5
Task 3	23.4	6:16	56.6	6.65	16.1

Table 1: Quantitative results of the user study

	Task 1	Task 2	Task 3	ALL
How well the tool helped to select?	3.9	3.1	4.1	3.7
How well the selected photos reflected the event?	4.5	3.9	4.6	4.3
Compared to regular way of selection	4.2	3.3	4.1	3.9

Table 2: User satisfaction results

ways of selecting photographs. The answers to these questions are summarised in Table 2. For each task the number of clicks and the time spent was measured, as well as the number of photos selected (see Table 1).

The short events the participants searched photos for were a birthday party, roller skating, and holiday trips. For the long events the participants all had a trip: hiking, traveling, and a long roller skating trip. For the yearbook task whole set of images was used and no temporal or event restrictions were given. The participants selected about 10-20 in each task to be sent to friends, family, or people who were in the photographs. In the case of the yearbook, the participants made the book mainly for themselves and planned to show it to friends and family.

The participants were satisfied on how well the photos they selected represented the event. In the long event task (task 2) they reported that they felt that they missed some photographs they would have liked to have. In the short event they felt that no photographs were missing, and in the yearbook task one participant reported that he got almost all of them, and another participant felt that she missed 5-6 photographs. As seen in Table 2, the participants were very happy with the photographs they had selected in tasks 1 and 3. In task 2 they thought they had missed some, but felt content anyway.

Overall, the FreeEye tool was scored high in our trials. As shown in Table 2, the overall average score for how well the tool helped the user in selecting photographs was 3.7 on a scale from 1-5 (1=very bad, 5=very good). Compared to the participants regular ways of selecting photographs for similar tasks it scored 3.9 on a scale of 1-5 where 3 was as good as their regular one and 5 was much better. All but one of the participants used Windows operating system's user interface to select their photographs, and the tool was considered better than Windows OS (average score of 4.1). The one participant used Picasa and he thought the tool was as good as Picasa (score of 3).

Generally the tool was thought to be good in recollecting events and photographs taken. The way in which it showed forgotten photographs was mentioned as a positive thing. One of the main issues the participants had with the tool was that if they had a particular photograph in their mind, it was not always easily found. Especially Task 2 (long event) was considered harder to do than the other tasks because there were more pictures than in a short event and unlike the yearbook task, the long event was restricted in time. The quantitative data in Table 1 supports this: more time was spent per chosen photograph than in the other tasks, although the time spent between clicks was not significantly different.

4.3 Cross-user photo selection and browsing

In this user study, the aim was to evaluate the bias present if the owner recalls the perceptual features of a specifically targeted photo when compared to a person who doesn't. This was achieved by cross-sharing of photos taken at the same event between two friends. Therefore, in our fourth experiment, we invited 5 couples of friends who own a significant collection of photos of the same event (on an average each participant contributed with 1500 photos). The task directions for both users were as follows:

- **Long**: choose a long event and photos from that event (longer than 2 days)
- **YearBook**: select photos for a yearbook for yourself
- **Gift**: select photos to give to a person as a gift on a CD-ROM/USB stick/website
- **Long (cross-user)**: from a long event that you participated, choose photos to share

In order to evaluate the user experience, users were asked to grade their responses to following questions:

Q1 On a scale of 1-5 (1=terrible, 5=excellent), how well did tool help you find what you wanted?

Q2 On a scale of 1-5 (1=terrible, 5=excellent), how well do the selected photos reflect the event? (for Long and YearBook tasks)

Q3 On a scale of 1-5 (1=terrible, 5=excellent), how good a gift the photos would be? (for Gift task)

Q4 On a scale of 1-5 (1=terrible, 5=excellent), how well did the tool help you in selecting photos for sharing?

Q5 On a scale of 1-5 (1=much worse, 5=much better), how would you rate this tool compared to what you normally use?

The quantitative results to user satisfaction are given in Table 3. It is observable that the overall user response was consistently positive with low deviation, when compared to their everyday practice. This is especially the case with the

Task	Q1	Q2/3	Q4	Q5
Long	4.0 ± 0.7	4.3 ± 0.7	3.9 ± 0.7	3.6 ± 0.8
	4.0 ± 0.7	4.3 ± 0.7	4.3 ± 0.5	3.6 ± 0.8
YearBook	4.0 ± 0.7	4.2 ± 0.8	4.2 ± 0.7	3.9 ± 0.7
	4.0 ± 0.7	4.3 ± 0.7	4.4 ± 0.6	4.1 ± 1.0
Gift	4.2 ± 0.6	4.4 ± 0.7	4.2 ± 0.6	3.8 ± 0.8
	4.0 ± 0.7	4.5 ± 0.7	4.2 ± 0.6	3.9 ± 0.9

Table 3: Cross-user satisfaction results

final outcome, since Q2 and Q3 scored very high results. The criticism was directed towards not being able to find a specific photo and a feeling of missing some events completely.

However, the quantitative analysis of user clicks and timings were inconsistent from user to user and from task to task, so that no conclusion could be drawn. This result is inline with the analysis of user-centred evaluation in interactive information retrieval [23]. This aspect of our study will be at the centre of our future studies in user-centric access to personal visual archives.

5. CONCLUSIONS

In this paper we have introduced a novel interactive interface for intuitive search and browsing of image. The presented interface is targeting a multitude of applications: from browsing of personal photo collections, find specific photos for short or long event to selecting a year photo book. From the experimental results, the system is very usable and intuitive, while offering pleasant browsing of visual data and often offering new perspectives of the same dataset by making surprising links between the data subsets. In addition, the users could manipulate the visual interface without any specific introduction. Finally, the knowledge discovery element of four random images in the corners of the display has been proven as a very useful tool of the interface.

Having in mind that our research interest is in building a user interface that leverages available information to facilitate the photo browsing, search and selection process, not to automate it, the results of the user studies are promising. The photo selection from increasingly large personal collections is a common task for a variety of situations. For that reason we have built a tool where the user is in charge and does the final selection. In our tool we used only the visual similarity information to help the user select photos for emailing, uploading, or making a book. Surprisingly, the visual similarity was considered helpful and as the scores of our trial show the participants were quite happy with the tool and the selected photographs. The evaluation outcomes can be summarised as follows:

- The selected photographs reflected the events very well (4.3/5)
- The tool was considered helpful (3.7/5), and better or as good as their existing ones (3.9/5)
- The participants selected on average 10-23 photographs, and spent from 2-6 minutes in selecting the photographs.

What we learned from our trial was that our tool seems to work well with personal collections: the participants knew their own photographs which helped them to feel in control. This became especially clear with one participant who had in her collection also photographs taken by someone else. This caused confusion and a feeling of being lost. The strength of our tool is that it is a general tool that is not coupled with any particular task or with any particular system. The other main strength is that according to our user trial, people found it useful and helpful.

6. REFERENCES

[1] M. Cooper, J. Foote, A. Girgensohn, and L. Wilcox. Temporal event clustering for digital photo collections. In *Proceedings of the eleventh ACM international conference on Multimedia*, pages 364–373, New York, NY, USA, 2003. ACM Press.

[2] A. Graham, H. Garcia-Molina, A. Paepcke, and T. Winograd. Time as essence for photo browsing through personal digital libraries. In *Proceedings of the Second ACM/IEEE-CS Joint Conference on Digital Libraries*, New York, NY, USA, 2002. ACM Press.

[3] A.Loui and A. E. Savakis. Automatic image event segmentation and quality screening for albuming applications. In *Proceedings of the IEEE International Conference on Multimedia and Expo*. IEEE, 2000.

[4] Kentaro Toyama, Ron Logan, and Asta Roseway. Geographic location tags on digital images. In *MULTIMEDIA '03: Proceedings of the eleventh ACM international conference on Multimedia*, pages 156–166, New York, NY, USA, 2003. ACM.

[5] Apple Ltd. *iPhoto'09*, 2009.

[6] David F. Huynh, Steven M. Drucker, Patrick Baudisch, and Curtis Wong. Time quilt: scaling up zoomable photo browsers for large, unstructured photo collections. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*, pages 1937–1940, New York, NY, USA, 2005. ACM.

[7] Benjamin B. Bederson. Photomesa: a zoomable image browser using quantum treemaps and bubblemaps. In *UIST '01: Proceedings of the 14th annual ACM symposium on User interface software and technology*, pages 71–80, New York, NY, USA, 2001. ACM.

[8] C G Jung. Psychological types. Princeton University Press, 1971.

[9] Di Zhong, Hongjiang Zhang, and Shih-Fu Chang. Clustering methods for video browsing and annotation. Technical report, In SPIE Conference on Storage and Retrieval for Image and Video Databases, 1997.

[10] Yueting Zhuang, Yong Rui, Thomas S. Huang, and Sharad Mehrotra. Adaptive key frame extraction using unsupervised clustering, 1998.

[11] Ahmet M. Ferman and A. Murat Tekalp. Multiscale content extraction and representation for video indexing. volume 3229, pages 23–31. SPIE, 1997.

[12] M Christel. Evaluation and user studies with respect to video summarization and browsing. *Proceedings of SPIE*, Jan 2006.

[13] J. Calic, D.P. Gibson, and N.W. Campbell. Efficient layout of comic-like video summaries. *Circuits and Systems for Video Technology, IEEE Transactions on*,

17(7):931–936, July 2007.

[14] Ngo Chong-Wah, Ma Yu-Fei, and Zhang Hong-Jiang. Video summarization and scene detection by graph modeling. *Circuits and Systems for Video Technology, IEEE Transactions on*, 15(2):296–305, 2005.

[15] Janko Ćalić and Neill W. Campbell. Compact visualisation of video summaries. *EURASIP J. Adv. Signal Process*, 2007(2):17–17, 2007.

[16] Pedro F. Felzenszwalb and Daniel P. Huttenlocher. Efficient graph-based image segmentation. *Int. J. Comput. Vision*, 59(2):167–181, 2004.

[17] David Kirk, Abigail Sellen, Carsten Rother, and Ken Wood. Understanding photowork. *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, April 2006.

[18] David Frohlich, Allan Kuchinsky, Celine Pering, Abbe Don, and Steven Ariss. Requirements for photoware. *CSCW '02: Proceedings of the 2002 ACM conference on Computer supported cooperative work*, November 2002.

[19] Lev Kuleshov and Ronald Levaco. *Kuleshov on film: writings by Lev Kuleshov*. University of California Press, Berkeley; London, 1974.

[20] Jacques Derrida. *Of grammatology*. Johns Hopkins University Press, Baltimore, London, 1997.

[21] Kan Ren and Janko Calic. Freeeye: interactive intuitive interface for large-scale image browsing. In *MM '09: Proceedings of the seventeen ACM international conference on Multimedia*, pages 757–760, New York, NY, USA, 2009. ACM.

[22] Kan Ren, Risto Sarvas, and Janko Calic. Freeeye: intuitive summarisation of photo collections. In *MM '09: Proceedings of the seventeen ACM international conference on Multimedia*, pages 1127–1128, New York, NY, USA, 2009. ACM.

[23] Daniela Petrelli. On the role of user-centred evaluation in the advancement of interactive information retrieval. *Information Processing & Management*, 44(1):22–38, 2008.

Description and Selection of Media Archives for Geographic Nearest Neighbor Queries in P2P Networks

Daniel Blank
University of Bamberg
D-96052 Bamberg, Germany
daniel.blank@uni-bamberg.de

Andreas Henrich
University of Bamberg
D-96052 Bamberg, Germany
andreas.henrich@uni-bamberg.de

ABSTRACT

In recent years, there has been a tremendous increase in personal media data stored on people's PCs, mobile devices, and social media sites in the web. Additionally, people are increasingly collaborating and interacting by sharing and commenting media items. These trends call for retrieval services integrating resources heterogeneous in update frequencies, media types, and size. In this context, peer-to-peer (P2P) technologies offer interesting solutions. When performing a query on certain types of P2P networks, resource selection is important. Compact summaries (i.e. resource descriptions) of each peer's data collection are known to other peers and used by them in order to determine promising peers for a given query. Summaries have to describe not only textual information, content-based media features, and information about date and time, but also the locations where e.g. images were taken, videos were recorded, or to which a user travelled. The present paper proposes and evaluates different resource selection techniques based on descriptions of the geographic footprint of personal media archives when querying for media items that are geographically close to a given query location. These techniques are not restricted to P2P networks and can e.g. be applied in hybrid index structures or distributed IR systems in general.

Categories and Subject Descriptors: H3.3 Information Storage and Retrieval: Information Search and Retrieval [Search process, Selection process]; H3.4 Information Storage and Retrieval: Systems and Software [Distributed systems, Performance evaluation]

General Terms: Algorithms, Performance

Keywords: Resource Selection, Peer-to-Peer IR, Multimedia IR, Summarization

1. INTRODUCTION

There has been a tremendous increase in personal media data during the last years. People write blogs, twitter about their lives, and use remote photo and video communities. Besides customized services such as Flickr or Youtube, cloud services offer more general storage facilities (e.g. `https://one.ubuntu.com/`, all URLs last visited on 13.1.10). In addition, personal information management (PIM) becomes important to administer e-mail accounts, personal contacts, appointments, tasks, and notes accessible from multiple devices. Consequently, heterogeneous online resources differing in size, media type, and update characteristics have to be administered [18]. Besides storing personal media items, people tend to share these items with friends and interact with each other by collaboratively tagging or commenting on various items.

In our scenario, personal media archives are administered in a P2P system. All media items of an archive are stored locally on the peer's/user's personal device without a need to store items on remote servers hosted by service providers such as Flickr or Youtube. Besides avoiding dependance on service providers as informational gatekeepers, no expensive infrastructure has to be maintained by applying a scalable P2P protocol such as Rumorama [20]. Idle computing power in times of inactivity can be used to maintain, analyze and enrich media items. For the purpose of sharing and collaboration, friends or groups can be granted access to certain media items of a peer/user.

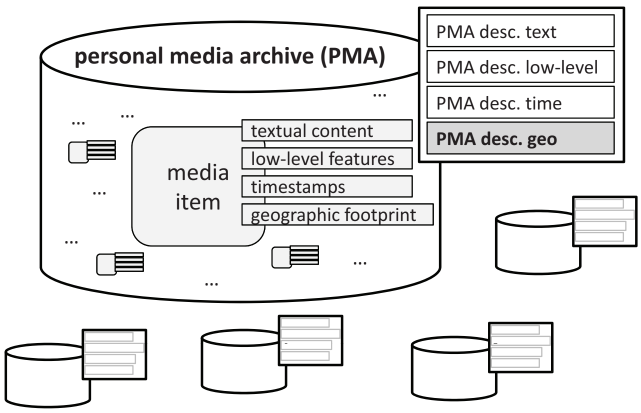


Figure 1: Criteria for media archive selection.

In order to facilitate retrieval, media items are described by four criteria: 1) textual content, 2) low-level features, 3) timestamps, and 4) a geographic footprint. Personal media archives containing multiple media items can thus be represented by four corresponding summaries (cf. Fig. 1) allowing for efficient and effective archive selection when processing a given query. In literature many approaches for P2P information retrieval (IR) can be found (cf. Sect. 2). However, these approaches mostly do not consider a unified search scenario combining all criteria from Fig. 1. To our best knowledge, the approach outlined here is the first with this target in mind. We analyze geographic resource description and selection techniques in this paper (highlighted in Fig. 1). Different geo-summaries are evaluated w.r.t. geographic nearest neighbor queries in Sect. 4. Techniques for summarizing textual and media content information in form of high-dimensional, real-valued feature vectors have already been proposed (cf. [6, 11] for example). Summarization of time and date information seems possible by a combination of clustering and histogram techniques (cf. Sect. 2). Of course, resource ranking based on a single criterion is only a first step. When querying for multiple criteria, individual resource rankings can be merged by applying a merging algorithm for ranked lists (cf. [3]).

Our approach is based on, but not restricted to, Rumorama [20], a scalable P2P protocol building hierarchies of PlanetP-like networks accessible by an efficient multicast. In PlanetP [11], randomized rumor spreading assures that every peer knows summaries of all other peers in the network. Summarizations of other peers' data provide the basis for query routing decisions, i.e. which peers to contact during query processing. While we examine peer summaries for geographic data in PlanetP-like middle-sized networks, results can be applied within large-scale Rumorama-like P2P networks. We also believe that results can be transferred to other application domains (cf. Sect. 2). Additionally, summarizations can be visualized (cf. Fig. 2 and 4). This might be beneficial for interactive retrieval e.g. by providing a visual overview of personal media archives for a huge number of archives with low bandwidth requirements.

The paper is organized as follows. Section 2 gives a brief overview on related work. In Section 3 we describe the resource description and selection techniques we employ. The experiments are part of Section 4. Here, we use two collections of geotagged images with user-information which allows for a realistic, user-oriented distribution of images to resources. Section 5 consists of a conclusion and an outlook on future work.

2. RELATED WORK

P2P IR systems can be classified into several groups. Systems such as PlanetP [11] and Rumorama [20] follow a semantic query routing approach based on resource descriptions generated and transmitted by every peer.

The second group of approaches are semantic overlay networks (e.g. [14]) where the content of the peer's data defines its place in the network topology. Peers are organized by semantic clusters and within query execution the query has to be forwarded to the most promising clusters. Here, the simultaneous indexing of multiple criteria as depicted in Fig. 1 would require the definition of a similarity between peers and images combining e.g. geographic and image content information. Alternatively, multiple overlays might be maintained.

A third class of P2P systems are structured networks such as distributed hashtables (DHTs). Novak et al. [7] present a large-scale architecture based on DHTs. Within DHTs, indexing data is transferred to remote peers with every peer being responsible for a certain range of the feature space. Presumably, correlations between different criteria (e.g. geographic and image content information) are difficult to exploit. If we e.g. assume an image from the Sahara Desert with shades of beige sand and blue sky, different peers might be responsible for indexing the geographic and the image content information. Therefore, when distributing the indexing data of the Sahara image, querying for it, or removing it from the network, (at least) two different peers have to be contacted. In the case of high-dimensional feature vectors, a problem within structured P2P networks is the load imposed onto the network when inserting new documents. If many peers do this at the same time, the insertion of data might become a bottleneck. Therefore, both, super-peer as well as summarization techniques are used within DHTs. While super-peer approaches are out of the scope of this paper, the strategy of creating peer summaries that are later indexed in a DHT is e.g. applied in [15] for content-based image features. In general, there is a convergence of structured and unstructured P2P networks with many hybrid approaches. DHTs have also been applied for social (semantic) desktops, e.g. in the Nepomuk project (`http://nepomuk.semanticdesktop.org`), which supports collaborative PIM and the sharing of media items. Nepomuk allows full-text search in combination with RDF-based queries. We see our work complementary in two directions. First, it tries to overcome some limitations of DHTs. Second, it allows for content-based multimedia retrieval enhanced with geographic nearest neighbor queries.

There is plenty of work on resource selection in traditional, distributed IR—especially for text data (for references see e.g. [18]). Most of the proposed resource descriptions are designed to maximize retrieval quality within the context of server selection. Within such a client/server scenario, space efficiency of resource descriptions is less important. Hence, summaries in distributed IR are usually less space efficient than the ones we are looking at.

Clustering (cf. [9]) might be used to represent collections of geo-locations by their centroids. Recently, cluster hulls have been proposed that summarize sets of geo-locations by several convex hulls [13]. We expect the summaries generated by these techniques to be less space efficient than the most promising approaches from Section 3. Here, image locations are assigned to the closest reference vector, which can be interpreted as a special form of clustering, leading to space efficient resource descriptions.

Tree-based index structures are also related to this work (cf. R-tree and variants [12, 16]). The decision of choosing the best subtree is similar to the peer selection problem. Summaries in the P2P context correspond to summaries maintained in the nodes of a tree (e.g. bounding boxes). Becker et al. [2] present an algorithm for summarizing a set of bounding boxes by two bounding boxes minimizing the area that is covered. Chen et al. [5] propose several threshold-based algorithms to split a single bounding box into several smaller ones in order to reduce the space within a bounding box that is not covered by any indexing data. These approaches demonstrate that there is optimization potential with the representation of geo-regions in spatial access structures. However the mentioned approaches stick with bounding boxes and the approaches presented in this paper might be interesting alternatives in this application field. A detailled comparison will be part of future work.

Our summarization techniques might also be applied within sensor [10] as well as ad hoc networks [15]. Because of limited processing power, bandwidth and energy capacities it is essential that aggregation techniques used within sensor networks are based on local information with a clear focus on space efficiency. Lupu et al. [15] present an interesting approach for ad hoc information sharing based on mobile devices when people meet at certain events or places. Here, it might not be feasible to share the complete data but only summarized information.

Compact resource descriptions might also be valuable for geographically focused crawling [1]. If a service provides summaries of the geographic extend covered by a certain website or media archive, a crawler could estimate the potential usefulness of this resource for its focused crawling task before actually visiting the source. This way, crawl efficiency can be improved by preventing the crawler from analyzing too many irrelevant pages. Geoparsing of web pages as well as downloading large sets of images in order to extract EXIF information can be avoided.

3. RESOURCE DESCRIPTION & SELECTION FOR GEOGRAPHIC QUERIES

The description and selection of resources is common in distributed IR (cf. Sect. 2). In general, there is a trade-off between the quality of resource descriptions and their size. Larger summaries can encode more information and should therefore allow for better resource selection. Within most of the distributed IR literature there is a clear focus on the quality of the resource descriptions. Often, only few resources with static data are analyzed making space efficiency of the resource descriptions not the main optimization target. We, on the contrary, need to find a more balanced solution. For our scenario it is essential to apply summaries that are space efficient and at the same time meaningful and selective enough to allow for efficient and effective resource selection, mainly because of two reasons. First, PlanetP-like networks might consist of more resources than many other systems in distributed IR. Second, resources might (re)join the network and the content administered by the resources might change frequently triggering new resource descriptions to be generated and sent.

Resource selection is performed by ranking the peers based on a set of resource descriptions, the query location and maybe some additional information such as reference points. The peer ranking defines the order in which peers are being contacted during query processing. When searching for the k closest images w.r.t. a query location, the peer ranking should reflect that peers administering a bigger fraction out of the top-k images receive a higher rank than peers maintaining a smaller fraction of top-k images.

In the following we will present four different resource description and corresponding selection techniques. The selectivity of three of the resource description techniques is briefly analyzed in [4]. Query processing is not discussed in [4]. Thus, the design and analysis of different resource selection techniques for k nearest neighbor queries and the improvement of the most promising resource description techniques are the main contribution of this paper.

Bounding Boxes (BB)

When using BB as resource description every peer computes a bounding box over the geographic coordinates of its image collection (cf. Fig. 2, left). We encode a latitude/longitude-pair (for short: lat/lon-pair) with 8 bytes, 4 for latitude and 4 for longitude. Therefore, we require $8*2$ bytes of raw data for the bounding box (i.e. two lat/lon-pairs, e.g. the lower left and upper right corner).

Peer ranking is performed as follows. If a peer p_a contains the query location within its bounding box whereas peer p_b does not, peer p_a is ranked higher than peer p_b and vice versa. In case the query location lies within the bounding box of both peers p_a and p_b, the size of a peer (i.e. the number of images a peer administers) is used as an additional criterion. Peers with more images are ranked higher than peers with fewer images. If neither peer p_a nor peer p_b contains the query within its resource description, the peer with the smaller minimum distance from the query location to its bounding box is preferred. We assume a spherical model of the earth with a radius of $6,371$ km. If not stated otherwise, we use Haversine distance [17] to compute the distance between two points on the sphere.

Grid-based Summaries ($GRID_r$)

In a second approach, the geographic coordinate space is represented as a grid (cf. Fig. 2, middle). A parameter r is used to define the number of rows of the grid. The number of columns is twice the number of rows since longitude range is twice as big as latitude range. The range of a grid cell (in degrees) is determined by $\frac{180^\circ}{r} = \frac{360^\circ}{2r}$ in the latitude and longitude domain. This simplified view is e.g. also applied in [8] and results in non-uniform grid cell sizes on the sphere. We gain selectivity and retrieval performance by increasing the number of grid cells at the price of additional storage overhead partially compensated through compression techniques (cf. Sect. 4.4). Every grid cell is represented by a single bit. If one or more image locations fall into a certain cell, the corresponding bit is set to 1. Otherwise, it remains 0. Within the summary, bit positions are determined horizontally from left to right and from bottom to top. Effects of alternative strategies on compression will be evaluated in future work.

During peer ranking the grid cell containing the query location is determined. If peer p_a has an image within this cell whereas peer p_b has not, peer p_a is ranked higher than peer p_b and vice versa. We also consider neighboring grid cells. If either both or none of peer p_a and peer p_b have an image located within the cell containing the query location, GRID considers the neighboring cells recursively until a ranking decision can be made. So, in the first round the ranking decision is always based on a single cell; in the second round it is in most cases[1] based on $1 + 8 = 9$ cells and in the third round on $1 + 8 + 16 = 25$ cells and so on. The ranking criterion in every round is the number of grid cells containing one or more image location(s) (the more the better).

Highly Fine-grained Summaries (HFS_n)

In this case we use resource descriptions originally designed for summarizing image content based on the color distribution or texture of an image [6]. We randomly choose a set of n image locations from the global image collection as

[1]This is not always the case since there might be no neighboring cells in a certain direction, e.g. as soon as a cell in the north or south is reached. Of course, at the 180-degree meridian we assume that there is no boundary and neighborhood relations are valid in both directions.

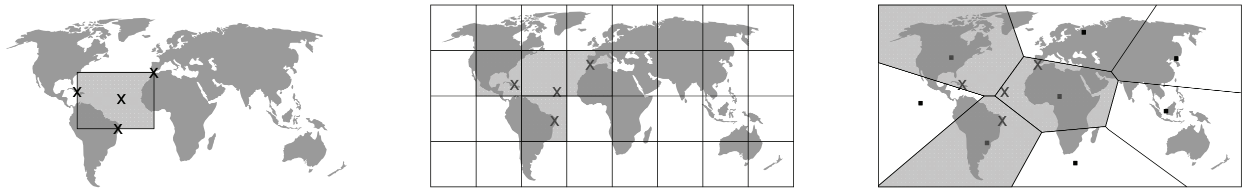


Figure 2: Visualizing summary creation for *BB* (left), *GRID*$_4$ (middle) and *HFS*$_8$/*UFS*$_8$ (right: ■ corresponds to reference points). Four images are geotagged, indicated as x.

reference points. This set of reference points is known to all peers[2]. Every image of a peer's local image collection is afterwards assigned to the closest reference point according to Haversine distance (cf. Fig. 2, right). Hereby, a cluster histogram is computed counting how many image locations of a peer's collection are closest to a certain reference point, i.e. cluster centroid c_j $(1 \leq j \leq n)$.

Peer ranking is performed as follows. The reference points c_j are sorted in ascending order according to Haversine distance w.r.t. the query. The first element of the sorted list L corresponds to the cluster centroid being closest to the query. Peers with more documents in this so called query cluster are ranked higher than peers with fewer documents in the query cluster. If two peers administer the same amount of documents in the analyzed cluster, the next element out of L is chosen and both peers are recursively ranked w.r.t. the number of documents within this cluster.

Ultra Fine-grained Summaries (UFS$_n$)

In contrast to HFS, UFS are based on a bit vector with the bit at position j indicating if centroid j is the closest centroid to one or more of a peer's image locations. Therefore, we obtain a bit vector of size n. Of course, there is some loss of information when switching from HFS to UFS with n staying constant. However, UFS have the potential of resulting in more space efficient resource descriptions. Potentially, this allows for more centroids being used which might result in similar or even improved retrieval performance compared to HFS. Among other aspects, this will be evaluated in Section 4.4.

4. EVALUATION

In the following we will present the data collections we use (Sect. 4.1), some rough calculations and basic considerations justifying the use of summarized geographic resource descriptions (Sect. 4.2), main characteristics of the experimental setting (Sect. 4.3) and the experiments themselves (Sect. 4.4).

4.1 Data Collections

We use two data collections of geotagged images:

Geoflickr: During 2007 we crawled a large amount of publicly available images which had been uploaded to Flickr. In our scenario every Flickr user operates a peer of its own. We therefore assign images to peers by means of the Flickr user ID. All of the crawled images are geotagged. After some data cleansing the Geoflickr collection consists of 406,450 geotagged images from 5,951 different users/peers.

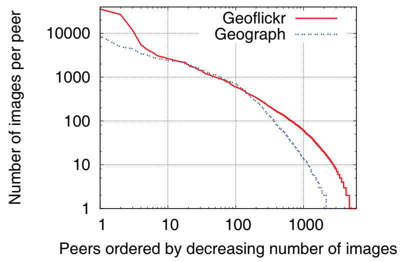


Figure 3: Number of images per peer for *Geoflickr* (5,951 peers) and *Geograph* (2,609 peers) collection.

Geograph: Geograph (`http://www.geograph.org.uk/`) "aims to collect geographically representative photographs and information for every square kilometre of Great Britain and Ireland". We downloaded the geotagged images and also distributed them to peers in a user-centric approach. In our scenario every Geograph participant operates a single peer; 2,609 peers administer 246,937 images in total.

The distribution of the number of images per peer is displayed for both collections in Figure 3. For both collections the distribution of the number of images per peer is skew which is typical for P2P networks [11]. Approximately the first 1% of the biggest peers per collection, i.e. the 60 biggest peers for the Geoflickr collection and 26 biggest peers for the Geograph collection, administer 42.0% and 28.3% of the collection's images respectively. Whereas the biggest peer of the Geoflickr collection maintains 8.8% of the images, the biggest peer of the Geograph collection maintains 3.5%. In opposition, approximately 20.7% and 17.7% of the peers administer only a single image for the Geoflickr and the Geograph collection respectively. Approximately 50% of the images are maintained by 1.8% resp. 2.7% of the peers for the Geoflickr and the Geograph collection.

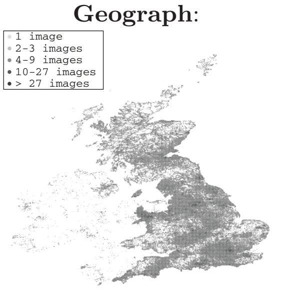
Figure 4 shows the geographic distribution of image locations. The Geoflickr collection consists of photos taken in various parts of the world with a focus on North America, Europe and Japan. In contrast, images of the Geograph collection are limited to the UK and Ireland with images more densily located around urban areas such as e.g. London.

4.2 Estimating Data Transfer Volume

The use of resource description and selection techniques is justifiable by several reasons. First, there might be scenarios e.g. in the context of ad hoc or sensor networks where it is not feasible or desirable to transfer complete indexing

[2]For CBIR, obtaining reference points from an external source and distributing them with software updates is proposed in [6]. We believe that this strategy is directly applicable also for geographic data. Strategies for selecting the reference points are evaluated in Sect. 4.4. Some peers could monitor the distribution of image locations in order to select appropriate reference points.

Geograph:



Geoflickr:

Figure 4: Geographic distribution of image locations for *Geograph* (top) and *Geoflickr* (bottom).

or sensor data (cf. Sect. 2). Second, resource descriptions might be used as "aggregators" to boost efficiency of certain types of applications such as focused crawling (cf. Sect. 2). Third, in certain scenarios the use of summaries will be beneficial compared to the transfer of full indexing data. Here, the total cost[3] of indexing (C^s_{idx}) and querying (C^s_q) information with the use of summarization techniques is smaller than the cost of transferring unsummarized indexing data C^u_{idx}. In the latter case resource selection becomes trivial, because the complete query processing can be performed locally ($C^u_q = 0$). So, resource selection is beneficial in terms of network traffic as long as:

$$C^s_{idx} + C^s_q < C^u_{idx}$$

We assume a message overhead m_o for header information, peer ID, timestamp, etc. and consider P resources and I documents/images. Furthermore, m_{idx} denotes the size of an index entry (e.g. 8 bytes in case of a lat/lon-pair) and m_s captures the average size of a resource description. Thus, in a non-dynamic network where resources stay online all the time and do not change their content, the cost of indexing is denoted as:

$$C^s_{idx} = (m_o + m_s) * P * (P - 1)$$
$$C^u_{idx} = m_o * P * (P - 1) + m_{idx} * I * (P - 1)$$

It follows that resource selection is beneficial as long as:

$$C^s_q < (P - 1) * (m_{idx} * I - m_s * P)$$

Thus, C^s_q has e.g. to be smaller than 15.8 GB and 4.5 GB for Geoflickr and Geograph respectively assuming average summary sizes of $m_s = 100$ bytes (cf. Sect. 4.4).

[3]In a rough calculation cost is measured in terms of data volume that is transferred and ignoring the storage requirements for replicating the indexing data.

An upper bound for the cost of querying summary-based networks can be derived depending on the average number of contacted peers $f * P$ ($0 \leq f \leq 1$), the cost for sending the query to a resource ($m_o + m_{idx}$) and the cost for receiving the result set ($m_o + m_{idx} * k$) where we assume that in the worst case w.r.t. C^s_q all inquired peers send k result items:

$$C^s_q = Q * f * P * (2m_o + m_{idx} * (k + 1))$$

For example, with UFS_{8192} on average 0.2% resp. 0.5% of peers are contacted for Geoflickr and Geograph (cf. Sect. 4.4). If we assume a message overhead of approximately 50 byte and Q queries in total, C^s_q becomes 3.2 $kB*Q$ and 1.4 $kB*Q$ for Geoflickr and Geograph indicating that under these conditions roughly 5.0 million and 1.3 million queries can be performed respectively, before exceeding the data volume C^u_{idx}.

Obviously, overall network load in a PlanetP-like setting also depends on the frequencies of peers (re)joining the network and the characteristics of peers updating their document collections. In general, the decision to use aggregated resource descriptions rather than the data itself has to be based on various network characteristics influenced by e.g. the application scenario and the type(s) of data being indexed.

Of course, there are further optimizations to our approach. For example, peers maintaining only few images might transfer the geo-locations directly whereas only big peers send aggregated resource descriptions. Such a hybrid resource description strategy affords hybrid peer ranking schemes which are not the scope of this work.

4.3 Experimental Setting

For every parameter combination we run at least 5,000 queries. If a randomized selection of reference points is needed we run 50 experiments with 100 queries each, which also results in 5,000 queries per parameter setting.

Space efficiency of different resource descriptions is measured by analyzing summary sizes (cf. Sect. 4.4). For compressing the summaries we apply Java's `gzip` implementation with default parameter values. Our measurements include serialization overhead necessary in order to distribute the resource descriptions within the network.

We use two modes for selecting a geo-location as query. First, we randomly choose a geo-location of an image from the entire document collection (*queryMode=1*). Since we do not remove the image with the query location, it is—on average—more likely that a big peer contributes to the retrieval result than a small peer because—on average—it is more likely to choose the query from a big peer than from a small peer. Second, we select a random peer and from this peer we choose a geo-location of an image at random (*queryMode=2*). Here, it is more likely that also a small peer contributes to the top-k query results since peers are chosen equiprobable.

When measuring retrieval performance we determine the fraction of peers that needs to be contacted on average in order to retrieve a certain fraction of the top-k image locations ($k = 20$) w.r.t. a given lat/lon-pair as query location. The top-k geo-locations are computed using Vincenty distance [19]. Since we are interested in the quality of the resource selection techniques, we analyze all of a peer's image locations as soon as it is contacted, because the top-k image locations of a peer determined using Haversine distance might differ from the top-k image locations computed using Vincenty distance. In a real-world application, only the top-k image locations will be transferred (together with some additional information such as peer ID, etc.).

4.4 Experiments

At first, we evaluate retrieval performance of the BB approach (cf. Sect. 3). Figure 5 (left) on the next page shows the fraction of peers contacted on average in order to retrieve a certain fraction of the 20 closest image locations to a given query location. For both collections it seems reasonable, in the case where the bounding box of both peers p_a and p_b contains the query location, to contact the peer that administers the larger number of images in total (*bc:size*). For *bc:bbsize* and *bc:bbrecipsize*, in the case when both bounding boxes of peers p_a and p_b contain the query location, an approximated size and the reciprocal of the approximated size of the bounding box is used respectively in order to make a peer ranking decision; *bc:bbsize* ranks a peer with larger approximated size of the bounding box higher while *bc:bbrecipsize* prefers a peer with smaller approximated size of the bounding box. For reasons of comparison, Fig. 5 (left) also includes a ranking solely based on the number of images a peer maintains (*size*). In this case, a peer is ranked higher the more images it administers.

For the Geoflickr collection *bc:bbrecipsize* performs similar to *bc:size* while preferring smaller peers since in general the approximated size of their bounding box is smaller compared to big peers. When solely ranking by peer *size* some big peers are contacted at an early stage that cannot contribute to the top-20.

It can be observed that for the Geograph collection the performance gap between *bc:size* on the one hand and *bc:bbsize* as well as *bc:bbrecipsize* on the other is more distinct compared to Geoflickr at least w.r.t. the retrieval of e.g. 80% of the top-k images, because in this case neither the approximated size of the bounding box nor the reciprocal are an adequate indicator for estimating the true *size* of a peer. In general, a ranking solely by peer *size* is less efficient for Geograph than for Geoflickr when querying for all of the top-20 image locations.

The size of a summary in the case of BB is 16 bytes for the bounding box plus 27 bytes serialization overhead, so 43 bytes in total. As it is desirable to include the size of a peer within the peer ranking process (*bc:size*), we assume additional 2 bytes for peer size information which leads to overall summary sizes of 45 bytes for BB (cf. Fig. 6).

BB performs better for the Geoflickr than for the Geograph collection which is likely to be due to the fact that for the latter there is more overlap among the bounding boxes [4]. Compared to the resource selection techniques presented in the following, BB performs worse. More peers are contacted on average in order to retrieve a certain fraction of the top-k query locations.

Retrieval performance of HFS/UFS is depicted in Fig. 5 (middle). As expected, differences in retrieval performance between HFS and UFS diminish with increasing number of summary bins n since the corresponding histograms become more and more similar with many zeros and some summary bin values set to 1. Of course, for HFS, the values of some summary bins might still be bigger than 1, but with increasing n this becomes rarer and rarer.

Using UFS seems worthwhile if in addition to retrieval

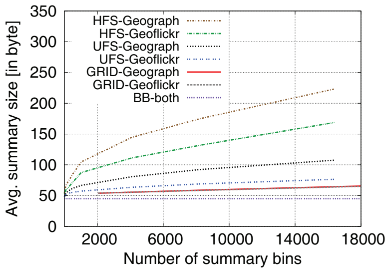


Figure 6: Avg. summary sizes, gzip applied to all but *BB*.

	Geoflickr HFS$_n$			Geograph HFS$_n$		
n	avg	min	max	avg	min	max
1,024	87.4	75.1	349.9	104.7	75.5	808.3
4,096	111.2	88.1	937.2	144.6	88.7	1,810.7
8,192	131.6	103.6	1,482.6	174.8	105.0	2,643.7
16,384	168.7	135.9	2,278.7	223.1	137.6	3,778.9
	Geoflickr UFS$_n$			**Geograph UFS$_n$**		
n	avg	min	max	avg	min	max
1,024	57.3	48.0	149.2	66.5	48.0	178.5
4,096	63.6	48.0	296.9	80.8	48.5	491.8
8,192	68.9	48.1	467.3	92.2	48.9	799.6
16,384	76.5	48.0	694.8	107.5	49.6	1,269.4

Table 1: Summary sizes for *HFS/UFS*

performance also overall summary sizes are analyzed (cf. Fig. 6). Average, minimum and maximum summary sizes are reduced when using UFS instead of HFS (cf. Tab. 1) indicating that both, resource descriptions of small as well as big peers can be represented in a more space efficient way. BB results in the most space efficient summaries, but retrieval performance is worse compared to HFS/UFS and GRID. Nevertheless, the use of BB might trade off when the network is not queried very frequently, which is a rather unrealistic assumption for P2P networks. GRID can be summarized more space efficiently than HFS and UFS, but retrieval performance is worse as will be shown in the following.

Results for GRID on both collections are shown in Fig. 5 (right). Obviously, retrieval performance is improved as more grid cells are used. For the Geograph collection, the grid is not adapted to the boundaries of the UK. Thus, retrieval performance is worse compared to the Geoflickr collection. Nevertheless, in the case where image locations are limited to a certain region of the world, an adaptation of the grid is possible. We did not adapt it in order to show the effects of a skew distribution of geographic image locations on a global scale. HFS and UFS are better suited for such scenarios than GRID, because they better adapt to the data that is used. Retrieval performance of HFS/UFS is also better compared to GRID when applied on a global scale, i.e. on the Geoflickr collection. Both, HFS_{8192}/UFS_{8192} as well as $GRID_{64}$ result in the same number of summary bins in the uncompressed case ($64 * 2 * 64 = 8192$). Assuming *queryMode=2* and the Geoflickr collection, in the case of HFS/UFS 0.2% of peers are contacted in order to retrieve all top-20 image locations, whereas in the case of GRID 2.3% of the peers are contacted on average.

So far, we have assumed that the n reference points are chosen from the underlying data collection. Although this approach is feasible in general, we will now evaluate different

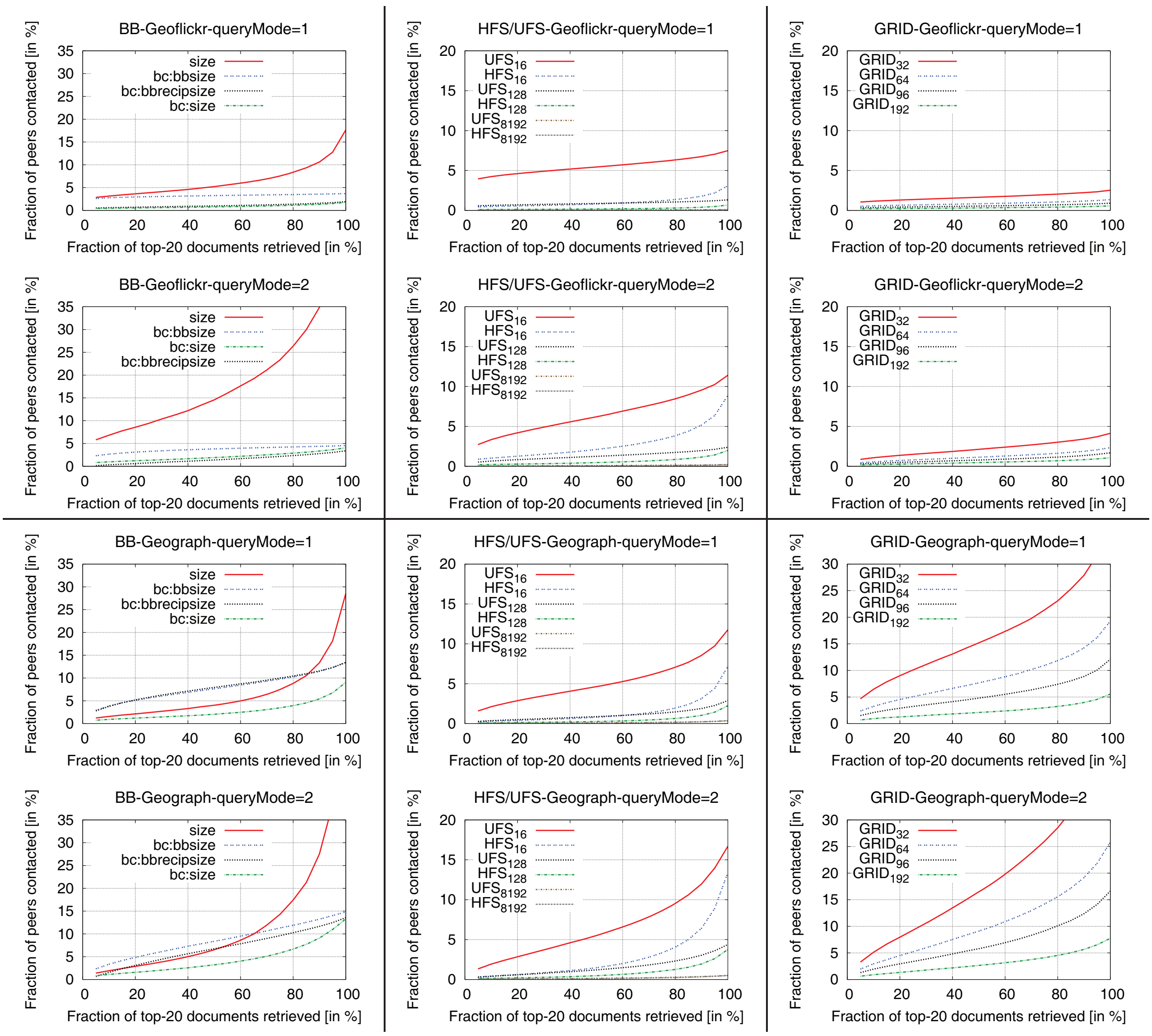


Figure 5: Fraction of peers contacted for *BB* (left), *HFS/UFS* (middle) and *GRID* (right)

sources for the reference points since we plan to distribute them with software updates from time to time in order to reduce overall network load.

We employ per country statistics from Worldmapper (`http://www.worldmapper.org/`) about mens' income (*INC*), Gross Domestic Product (*GDP*), population (*POP*), WWW usage (*WWW*) and land area (*LAND*). Based on these statistics we proportionally select the number of reference points from a certain country using Geonames gazetteer (`http://www.geonames.org/`). Reference points are selected amongst all populated places of a certain country at random. So, for example, if x% of the world's total mens' income is earned in a certain country, x% of the reference points are randomly chosen amongst all populated places of the specific country according to the information given by the Geonames gazetteer. For comparison we also choose reference points randomly distributed on the sphere (*RAND*).

Additional information is used from World Gazetteer (`http://world-gazetteer.com/`). We extract lists containing the geo-locations of the biggest cities in the *UK*, Europe (*EUR*) and the World (*GLOB*) in terms of population. A set of n reference points is chosen from each of these lists containing the n geographic locations with the most inhabitants according to the specific region.

Figure 7 shows results for *queryMode=2*. For reasons of brevity, we do not include figures for *queryMode=1* since in general they offer better retrieval performance as noticed before (cf. Fig. 5) and the same relative behavior as depicted for different strategies in Fig. 7. *COLL* represents the strategy where reference points are not determined from external sources but randomly chosen from the underlying data collection. On a global scale, when using external sources, selecting the centroids according to GDP is the most promising approach with performance similar to WWW and INC. These techniques adapt best to the data collections that are used and perform better than selecting e.g. the biggest cities in the world (GLOB). For both collections retrieval performance can be improved by increasing k e.g. up to $8,192$ or even higher. In general, Fig. 7 shows that *average* retrieval performance can be optimized by adequately selecting the

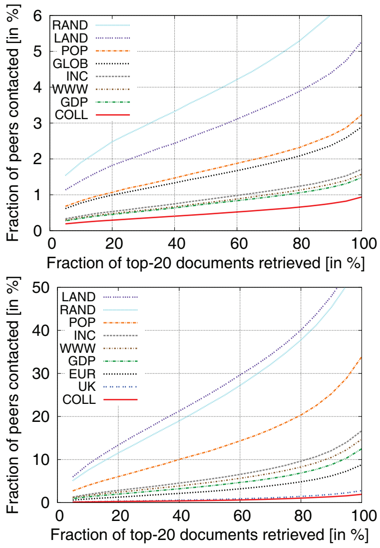


Figure 7: Sources of centroids for UFS_{512}; *Geoflickr* (top) and *Geograph* (bottom)

centroids according to the expected origin of query locations as well as image locations that are administered. This can e.g. be done by the implantation of special peers that analyze queries and the distribution of image locations. In Fig. 7 (bottom) the selection of centroids adapts to the collection by using the 512 biggest cities in the UK.

5. CONCLUSION & OUTLOOK

We have motivated a P2P system based on the description and selection of personal media archives. Media items are described by textual content, low-level multimedia features, timestamps, and geographic footprints. The focus of this paper is resource selection based on geographic information. In our experiments, both, bounding boxes as well as grid-based representations are outperformed in terms of retrieval performance by an approach using a set of reference points in order to compute (binary) histograms. Grid-based summaries show a higher potential for compression which might justify their use as well—depending on the characteristics of the usage scenario. Bounding boxes might be too coarse and induce overlap amongst peer summaries making selective peer ranking decisions difficult. Future work will mainly address the adaptation of UFS within CBIR as well as stopping criteria in order to determine when it is no longer beneficial to contact further peers during query processing.

Acknowledgement: Our work is partially funded by the Deutsche Forschungsgemeinschaft DFG HE 2555/12-2. We acknowledge the Geograph contributors (`http://www.geograph.org.uk/credits`). Their work is available under a Creative Commons Attribution-ShareAlike 2.0 Licence: `http://creativecommons.org/licenses/by-sa/2.0/`.

6. REFERENCES

[1] D. Ahlers and S. Boll. Adaptive geospatially focused crawling. In *Int. Conf. on Information and Knowledge Management*, pages 445–454, Hong Kong, China, 2009. ACM.

[2] B. Becker, P. G. Franciosa, S. Gschwind, T. Ohler, G. Thiemt, and P. Widmayer. An optimal algorithm for approximating a set of rectangles by two minimum area rectangles. In *Int. Workshop on Computational Geometry*, volume 553 of *LNCS*, pages 13–25. Springer, 1991.

[3] N. J. Belkin, P. Kantor, E. A. Fox, and J. A. Shaw. Combining the evidence of multiple query representations for information retrieval. *Inf. Process. Manage.*, 31(3):431–448, 1995.

[4] D. Blank and A. Henrich. Summarizing georeferenced photo collections for image retrieval in P2P networks. In *Proc. of Workshop on Geographic Information on the Internet*, pages 55–60, `http://georama-project.labs.exalead.com/workshop/GIIW-proceedings.pdf` (12.11.2009), Toulouse, France, 2009.

[5] Y.-Y. Chen, T. Suel, and A. Markowetz. Efficient query processing in geographic web search engines. In *SIGMOD Conf.*, pages 277–288, Chicago, IL, USA, 2006. ACM.

[6] D. Blank, S. El Allali, W. Müller, A. Henrich. Sample-based creation of peer summaries for efficient similarity search in scalable peer-to-peer networks. *ACM SIGMM Workshop on Multimedia Information Retrieval (MIR 2007), Augsburg, Germany*, pages 143–152, 2007.

[7] D. Novak, M. Batko, P. Zezula. Web-scale system for image similarity search: When the dreams are coming true. In *Int. Workshop on Content-Based Multimedia Indexing*, pages 446–453, London, UK, 2008. IEEE.

[8] R. Dolin, D. Agrawal, A. E. Abbadi, and L. K. Dillon. Pharos: A scalable distributed architecture for locating heterogeneous information sources. In *Int. Conf. on Information and Knowledge Management*, pages 348–355, Las Vegas, Nevada, 1997. ACM.

[9] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley-Interscience, 2 edition, Nov. 2000.

[10] B. M. Elahi, K. Römer, B. Ostermaier, M. Fahrmair, and W. Kellerer. Sensor ranking: A primitive for efficient content-based sensor search. In *Int. Conf. on Information Processing in Sensor Networks*, pages 217–228, Washington, DC, USA, 2009. IEEE.

[11] F. Cuenca-Acuna and C. Peery and R.P. Martin and T.D. Nguyen. PlanetP: Using gossiping to build content addressable peer-to-peer information sharing communities. In *IEEE Int. Symp. on High Performance Distributed Computing*, pages 236–246, Seattle, WA, USA, 2003.

[12] A. Guttman. R-trees: A dynamic index structure for spatial searching. In *SIGMOD Conf.*, pages 47–57, Boston, MA, 1984. ACM.

[13] J. Hershberger and N. Shrivastava and S. Suri. Summarizing spatial data streams using clusterhulls. *ACM Journal of Experimental Algorithmics*, 13:2.4–2.28, 2008.

[14] I. King, C. H. Ng, and K. C. Sia. Distributed content-based visual information retrieval system on peer-to-peer networks. *ACM Trans. Inf. Syst.*, 22(3):477–501, 2004.

[15] M. Lupu, J. Li, B. C. Ooi, and S. Shi. Clustering wavelets to speed-up data dissemination in structured P2P manets. In *ICDE*, pages 386–395, Istanbul, Turkey, 2007. IEEE.

[16] H. Samet. *Foundations of Multidimensional and Metric Data Structures*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2006.

[17] R. Sinnott. Virtues of the haversine. *Sky and Telescope*, 68(2):158, 1984.

[18] P. Thomas and D. Hawking. Server selection methods in personal metasearch: a comparative empirical study. *Information Retrieval*, 12(5):581–604, 2009.

[19] T. Vincenty. Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations. *Survey Review*, 22(176):88–93, 1975.

[20] W. Müller and M. Eisenhardt and A. Henrich. Scalable summary based retrieval in P2P networks. In *Int. Conf. on Information and Knowledge Management*, pages 586–593, Bremen, Germany, 2005. ACM.

Search and navigation as retrieval strategies in large photo collections

Ivar Solheim, Ingvar Tjøstheim, Øystein Dale, Lothar Fritsch, Till Halbach, Knut Holmqvist.
<name1.name2>@nr.no
Norwegian Computing Center
PO Box 114 Blindern
0314 Oslo Norway
Tel: +47 22852590

ABSTRACT
In this exploratory study we investigated the use of search and navigation as strategies for retrieval in large collections of digital photos. The main goal of the research was to test out models and methods that can describe retrieval behaviour and preferences. A focus group interview was conducted and demonstrated the need for taking various types of factors and measurements into account. We examined relationships between independent variables (happiness levels, satisfaction with and confidence in the search results, feeling lost during search,), and perceived satisfaction as dependent variable. The analysis showed statistical significant relationships for some important factors, but also indicated limitations in the applied theoretical framework, in particular related to use of emotions as independent variables. The study also showed that users' own perception of which retrieval strategy was the fastest differed from the actual time they used on search and navigation tasks. This confirms the need for analytical models that integrates subjective preferences/perceptions and measurement of objective factors.

Categories and Subject Descriptors
H5.2 User Interfaces: Evaluation/methodology; H3.3 Information Search and Retrieval

General Terms
Experimentation, Human Factors, Measurement

1. INTRODUCTION
This paper is a result from work in the Mariage (NCC 2009) research project, which stands for *Making Rich Media Accessible for Generations*. The project aimed at the development of principles, frameworks and demonstrators for life-time personal multi-medial albums. Media types of interest are photographs, videos, music and software-based media such as web pages, flash films/animations, and computer games.

This article addresses the aspect of *understanding strategies for end user retrieval of digital photos in large photo collections*. In the Mariage perspective, the development of sustainable and user friendly strategies for retrieval in multimedia collections is important.

The research question is to study users' preferences for two different retrieval strategies, *search* and *navigation*. *Search* means using the text search function, by applying keywords that matches photo metadata, e.g time and place. *Navigation* means to navigate visually through "tree" structures and pictures in photo collections, following implicit structures generated from metadata.

The experimental setting was designed to enable measuring of participants' subjective feelings before and immediately after search and navigation sessions. Furthermore, the study examines the relationship between the subjective factors and objective behavioural factors, such as time used on different tasks.

In this paper, "subjective factors" refers to the participants' own reported feelings, perceptions and assessments (Gwizdka 2007) before, during and after the task.

Previous related studies have reported lack of significant relationship between mood prior to the search/navigation, task difficulty level, searcher's interest in a task and positive and negative affects. This reflects that the relation between search and mood is a complex one from a research perspective, in particular in terms of methodically sound design (Nahl and Bilal 2004). We believe that some of the methodological challenges are partly due to the design and operationalization of the variables involved. One research aim of this explorative study is to provide improvements of how the variables are operationalized and formulated in the questionnaire to be used by the participants.

After a brief review of previous studies, we present our model and approach. This is followed by a description of methods, results from a preparatory focus group, the participants, tasks and procedure. The article then presents and discusses the results.

2. Related research
Our research approach departs from the basic assumption in the literature (Lopotovska 2009, Gwidzka and Spence 2007, Kules and Shneiderman 2008) that search and navigation strategies, and how users perform in carrying out such tasks, are influenced by subjective factors such as feelings, perceptions and evaluations

The present study extends the line of previous research in three ways:

(1) by focusing specifically on the study of search as well as navigation in digital multi-media material,

(3) focus on improving the operationalization of central variables (in particular the independent variables).

(4) a broader empirical approach, comprising both qualitative and quantitative approaches and methods.

Recent studies of search and navigation in textual material seem to indicate that the most efficient and user friendly approach is to *combine* search and navigation strategies in what is called a *faceted search* approach. (Li and Belkin 2008) However, research is less conclusive concerning which strategy will be preferred by users in retrieving multimodal material, such as photo collections. We believe it may be fruitful to study search and navigation experimentally in a photo collection setting.

The literature provides two ways to describe subjective states. First, ask participants *how they feel*. Second, ask participants how *happy they feel*. Both approaches have its merits and are widely applied in research, the first dominates in information science and market research, the second in measuring well-being in psychology (Kahneman, 2000).

The study draws upon previous studies (see Gwizdka and Spence 2007 for an overview) that have examined the relationships between searchers' subjective states and their behavior. Much of the work has been carried out in information science, but also studies in psychology shed light on several relevant aspects related to particular subjective factors, e.g. the notion of happiness (Kahneman, 2000)

Subjective aspects of information searching include aspects pertaining to the user's perceptions and feelings. The literature in the field includes studies of searchers' satisfaction (Su, 2003), relevance judgments (Saracevic, 2007) and feelings associated with the search stages (Marchionini, 1995; Agosto, 2002).

In previous inquiries in search studies, the approach to describing subject states is to ask participants how they feel. Various types of feelings can affect searcher's performance, but the feelings can also be affected by various other factors, such as user interface and the difficulty of the task.

There are several methodical challenges that must be dealt with in this type of research. Previous related studies (Lopatovska 2009a and b) have reported that due to several reasons there seems to be a lacking statistically significant relationships between mood prior to the search, topic of the search, sequence of task, task difficulty level, searcher's interest in a task and positive and negative affects (Lopatovska 2009). The present study will take into account that measuring mood in an experimental setting is often demanding. Clearly, search is a highly complex task for research (Nahl & Bilal, 2007) and participants' mood in the experimental setting may be affected by a number of factors that are not related to the task. However, lack of significant results may also be related to how the main themes, question and items are operationalized and formulated.

3. RESULTS FROM FOCUS GROUP

A focus group interview (Lindlof and Taylor 2002) was conducted to specify the research questions, increase our understanding of actual patterns of use and to provide input for the design of the planned experiment. It was important to gain insights about user strategies and behaviour from actual users.

The focus group comprised six people, varying in age from 20 to 73. All had experience and interest in photos and photography, but only one could be labelled an expert user. All participants were active photographers, most of them on a daily basis. The researchers had prepared an interview guide and had defined several themes for the group's work: equipment, hardware and software used, organization of collections, tagging/metadata, sharing with others, security, privacy and retrieval practices and experiences. The focus group interview lasted about two hours. The participants argued that navigation provided a feeling of closeness and personal control of the photos that a search alternative would not provide in the same way.

A striking result was that although all participants had large photo collections (only the expert user had systematically tagged his photos), it was not meaningful for them to apply search as the main method for retrieval. Time was the main organising factor, (typically a folder for each month and the subfolders for week and also days, and they typically used the options (year, month, date) offered in Windows Photo Gallery. The participants argued that navigation provided a stronger feeling of meaning, closeness and personal control of the photos than a search alternative would provide.

Another result relevant in the Mariage context, was that the young and the middle aged participants focused on creating *albums,* i.e carefully selected photos (perhaps 1 to 100) that were gathered not only for personal use, but also to be made available for friends and family. Sharing photos was a central goal. According to the participants' preferences, *socia*l considerations seem to play a more central role than what was expected by the researchers. For the younger participants, making sense of the photo collection depended heavily on how albums could be seen by others. For these participants, their photo collection was seen as raw material for the construction of physical albums - which can be seen as the ultimate goal or end-product for their photographing endeavours. The participants (in particular the younger ones) clearly favoured to construct and distribute a physical (rather than digital) album. Several of the participants used such albums as Christmas gift to friends and family.

The way the younger participants used Facebook was also interesting. They used Facebook as a social medium just as most others use it. They also presented personal photos like others do. What was unexpected and interesting was that they also used Facebook as one of the main storage medium for their own photo collections.

The focus group participants were all dedicated amateur photographers with large collections of photos that require a systematic approach to management and maintenance tools and practices. It is a paradox that the participants applied storage and search mechanism that were surprisingly simple and non-efficient in terms of time and effort spent. The participants applied an incremental "bottom-up approach" in managing and searching their photo collections. They preferred navigation rather than search even though this was a more time-consuming and less efficient approach. No metadata except the ones automatically produced by the camera were applied to the material. However, it can be speculated that users would have preferred search if their photos had been supplied with proper metadata at the outset, i.e. in the initial process of designing and constructing the photo collection structure. To apply metadata *after the fact* is more demanding, time consuming and less motivating; there may be thousands of photos that need to be post-processed.

The focus group seemed to indicate that the navigation mode has a strong position due to users' preferences for subjective control and overview in order to be able to select the relevant and meaningful photos. However, their preference for navigation can also be interpreted as a *fait accompli* since most participants have photo collections that are not designed for and equipped with relevant metadata that may make search a viable and realistic option. Search is simply not a relevant alternative. Thus, a preference for navigation can be expected. A fair hypothesis may be that user might prefer search if they know exactly what they are looking for and the photo database has metadata/tags that enable detailed search.

4. METHODS

In the following section the conceptual and methodological basis for the experiment is presented. A brief presentation of concept and how they are operationalized is followed by descriptions of the participants, the image collection and the task design and tasks.

4.1 Measuring feelings, perceptions and moods

4.1.1 Perceptions and feelings

Participants filled out questionnaires after each task. These questionnaires comprised two main sections, one section about perceptions and feelings and another section about mood. The section about perception and feelings such as

- Perceived level of *difficulty*,
- *Feeling* while searching
- *Getting lost*
- *Confidence*.

In order to increase level of reliability, these variables are seen as *constructs* of three defining, similar questions/statements in terms of content, but with changes in language and wording. For example, the questions defining and operationalizing Level of difficulty are:

1. Was it easy or difficult to solve this task
2. It is complicated to carry out the task
3 It was difficult to find the pictures that I was asked to find.

The participants shall fill out their response on a scale of 1 to 7

The operationalization of the variable constructs was an important research effort in its own right.

4.1.2 Mood

Participants' mood was measured using Positive Affect (PA) and Negative Affect (NA) Schedule (PANAS). The PANAS comprised of two 10-item scales that measure positive affect (feeling enthusiastic, alert, active etc.) and negative affect (feeling of anger, afraid, guilt, nervousness, etc.). A typical way of using PANAS is to measure past and present moods (Mackinnon, Jorm, Christensen, Korten, Jacomb, & Rodgers, 1999). In this project the PANAS form was filled out after the completion of each of the six tasks. PA and NA questions presented to participants after search tasks asked for their feelings just after the completion of the tasks.

4.2 Participants

Twenty subjects participated in a study conducted in a controlled experimental setting. Participants were students recruited from University of Oslo and Oslo University College. The inclusion criteria were simply that they spoke and read Norwegian well and were capable ICT users (not experts). They ranged in age from 20 to 25 years, a few older, but all less than 30. Participants were offered a monetary incentive (300 NOK, approximately 35€).

4.3 Photo collection and user interfaces

The photo collection used in our experiments was constructed by compiling a subset of photos from a freely available collection of photos of an existing family that has made several thousands of their own photos public under a Creative Common License[1] through Flickr[2]. The collection used in the test comprised 1000 photos that were selected from the larger collection. Each original photo had been tagged, i.e., described with a few words, by the owner, and this description was attached as metadata to the respective image. The photos were typically described with a name (*who*), a couple of words describing what is going on (*what*), and some information about the occasion (*where and when*).

The photos were made available for the participants through Picasa, an image management tool offered by Google. We chose Picasa 3.0 as it allows to conduct both image navigation and search, it has support for all relevant meta tags, and it has a simple, intuitive and very responsive user interface. Moreover, it is freely available on all major platforms, namely Windows, Mac, and Linux. Other interfaces and tools were considered, e g. Windows and iPhoto, but Picasa was chosen mainly because it was less intrusive and more transparent and intuitive for users and thus provided an efficient and sustainable tool for studies of search and navigation from the user's perspective. All person names are changed except for the original labeling and other relevant metadata such as time.

The user interfaces for both search and navigation are shown below in Figure 1 and 2 respectively.

---

[1] http://creativecommons.org/licenses/by-nc-sa/2.0/deed.en

[2] www.flickr.com

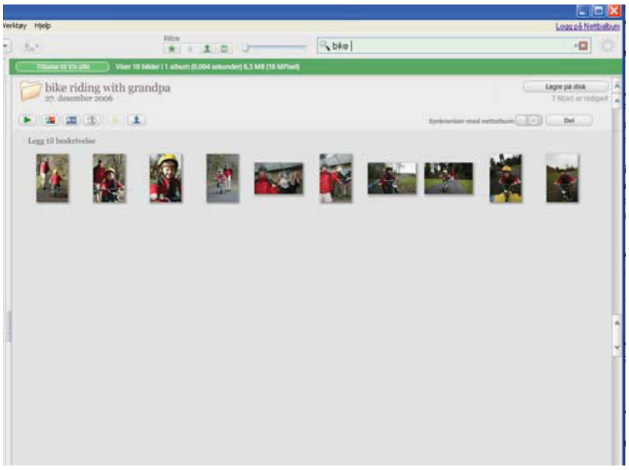


Figure 1 Search mode in Picasa.

The figure above shows the search mode. The participant is supposed to use the search field. Relevant input (e.g "bike") in this field will immediately (one letter is enough) show relevant photos (involving "bike"). The navigation structure is not visible.

In the figure below the navigation mode is illustrated. The photos are on the right hand side, the navigation structure is on the left hand side.

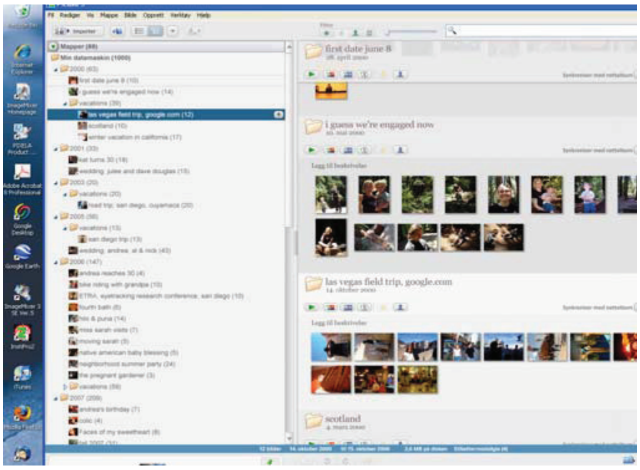


Figure 2 Navigation Mode in Picasa.

4.4 Task design and procedure

The experiment used a 1x2 within-subject (Kules and Capra 2008) design in which all participants used the same interface to complete two types of tasks, *search* type and *navigation* type search, alternating the mode used first. 10 participants started with search tasks, 10 with the navigation tasks

4.4.1 User Scenario and Tasks

The participants were presented with the following scenario.

"*You are supposed to assist the Smith family in retrieving pictures that shall be put into two new family albums about the Smiths. The family intends to give the albums to friends and relatives. The family has specific wishes about what pictures to select, but they need your assistance in retrieving these. You shall carry out six different tasks that will lead you to the preferred pictures.*"

Search tasks

1a. Search for pictures of Al, Andrea, Nick and Paul from the two recent years (2008 and 2009) and put them in the album. Find two pictures of each person, or alternatively pictures that show several of them together.

2a. Search for pictures from all trips to San Diego and find one picture from each trip in which members of the family have participated

3a. Search for pictures from the three most recent years that show Nick's and/or Paul's sporting activities, soccer and cycling.

Navigation tasks

1b. Use navigation to retrieve pictures of Al, Andrea, Nick and Paul from the two recent years 2008 and 2009. Find two pictures of each person, or alternatively pictures that show several of them together.

2b. Navigate to find pictures from all trips to San Diego and find one picture from each trip in which members of the family have participated.

3b. Navigate to find pictures from the three most recent years that show Nick's and/or Paul's sporting activities, soccer and cycling.

4.4.2 Setting and procedure

An experiment took place in a laboratory with a laptop computer , and was filmed with a video camera. The computer recording software Morae[3] recorded the session.

Each participant was scheduled for an individual session lasting from 50 to 120 minutes. The procedure and tasks to be carried out by the participants were:

- Upon arrival read and sign the consent form and listen to the explanation of the procedure.
- Fill out a pre-task questionnaire.
- Familiarization with the Picasa software, with an example to demonstrate the interface and the functionality.
- Read the scenario and the tasks
- Conduct search/navigation task (six tasks)
- Fill out questionnaire after each task
- Fill out post-task questionnaire and the open-ended questions about search versus navigation
- Short Interview concerning the answers to the open-ended questions and any other matter
- Receive NOK 300.

[3] Morae by TechSmith is a software that records real-world actions, such as user speech and facial expressions, along with detailed application and computer system data to provide a view into the way that web sites and software are seen and experienced.

5. Results from experiment

The findings presented in this section serve two purposes. First, we show some illustrative results based on the collected data. Second, we discuss and show some possibilities as well as limitations of the analytical and theoretical models that are applied.

Partial least square (PLS) is the statistical analysis technique applied to interpret how the factors are associated with each other. PLS[4] can be used with small sample, but there is a limitation related to the number of factors in the structural model (Chin, 1998). In this study five concepts are used as independent variables. These are level of difficulty, level of confidence, to what extent the user got lost, negative and positive emotions. The sample size indicates that a structural model should contain two of these concepts and one dependent variable. Because of the second purpose, all five concepts are included in the model presented below.

In a structural model, it is possible to analyze to what degree, if any, these concepts are associated with or have an influence on dependent variables such as satisfaction and use of time. Use of time is an efficiency measurement which can be measured by the Morae software that was applied in this study.

The figures 3 and 4 present the comparison of use of time for search vs. navigation. In both groups task 1, 2 and 3 were performed. The only difference is the mode. Figure 1 presents the comparison of search vs. navigation when the participants did the task for the first time. After doing the search, the same person did navigation. Hence, the tasks 1, 2 and 3 were repeated. The comparison of use of time in the repetition is shown in figure 4. The scale is on interval level. 1 means up to 119 seconds, 2 means 120 to 179 seconds, 3 means 180 to 239 seconds etc.

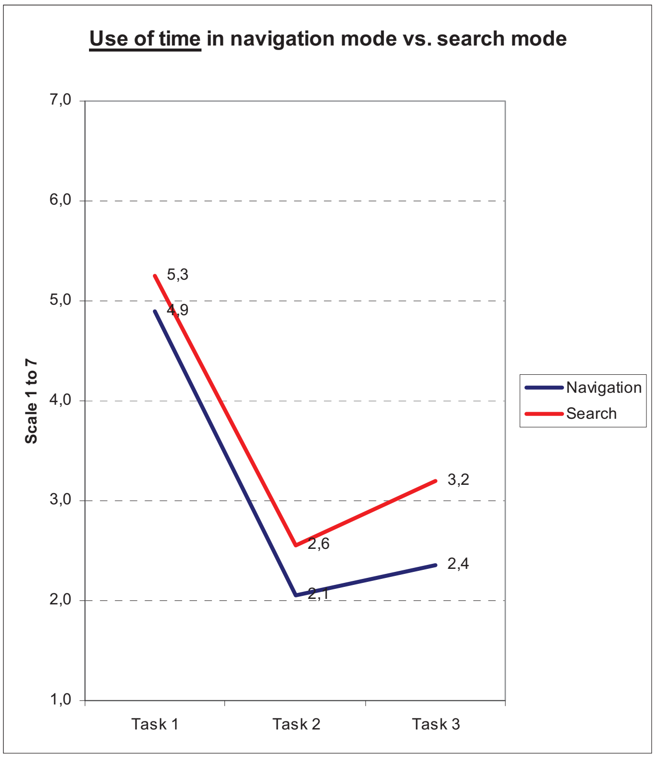


Figure 3 Comparison of use of time between navigation mode and search mode

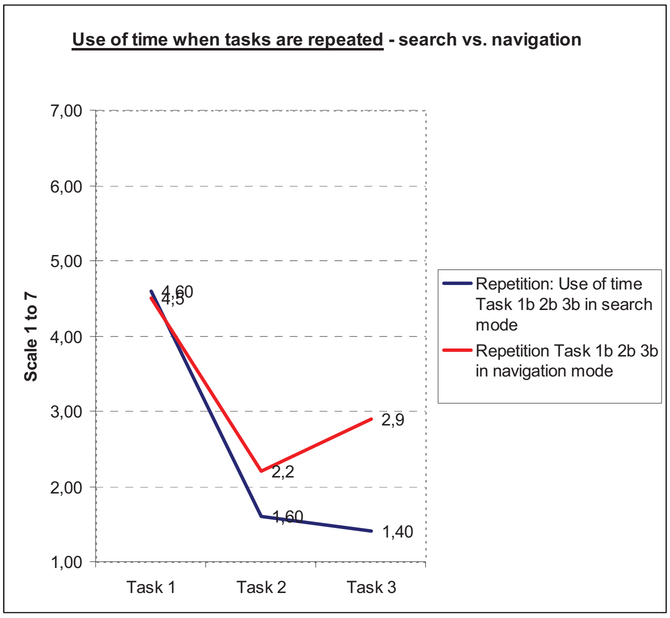


Figure 4 Comparison of use of time between navigation mode and search mode when the tasks are performed a second time.

For both modes (search and navigation) the respondents used less time on the tasks when the tasks were repeated. The findings indicate that the navigation mode is the faster of the two modes. In their comments 12 of the 20 respondents wrote that search is faster than navigation. Since the tasks are time-stamped, it is possible to judge this subjective assessment of the users. In both groups 6 out of 10 stated that search is faster than navigation. For the group that did navigation first, search was the fastest mode for all of them. However, this is primarily due to the learning effect. In the group that did search first and then navigation, 6 of 10 also stated that search is the fastest mode, but for 5 of the 6 this was not the case. The findings indicate that users are not accurate when estimating how much time they use on a task and which mode that is the faster. The pattern revealed is a reduction in time spent on the tasks when the tasks are repeated. A reasonable explanation for this is the learning effect.

The users were also asked about which mode they preferred. Although a majority (11) stated in their free comments that they believe search was the faster, 9 of 20 informants preferred navigation when asked to choose between modes. The informants seemed to appreciate other virtues than speed. One informant formulated what seemed to be a common view: "It is better to search when you are going to find specific photos. Navigation can be ok if you would like an overview of all the photos you have got".

A model that depicts the relationships between factors and concepts has to be based on former studies in the field. The model (see below, Figure 5) includes factors used by Kules & Capra (2008), Gwizdka & Spence (2006), Lopatovska (2009a and b). Also the PANAS framework is included here.

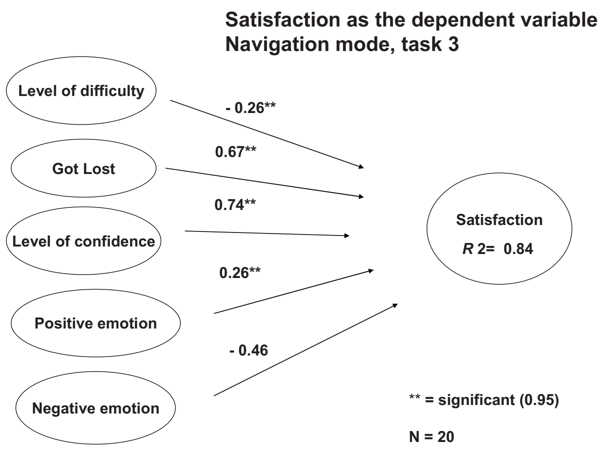


Figure 5; Structural model and results

The results can be interpreted as follows. The R^2 is on a high level. All five factors are significant and contribute in explaining the variance of the dependent variable satisfaction. Hence, the model explains a substantial amount of variation of satisfaction for task 3. For the factor getting lost "disagree" is computed as a low number and "agree" as a high number. Therefore, when the participants reported getting lost while doing the task, this has a negative effect on satisfaction. Positive emotions have a positive effect on satisfaction. This is also the case for level of confidence. Level of difficulty and negative emotions are also significant and contribute to explaining the amount of variance in the model.

A composite reliability metric for internal consistency assesses construct reliability. Table 1 below shows that all construct reliabilities exceed Nunnally's (1978) suggested 0.7 benchmark. The alphas are all above 0.6 which is regarded as a benchmark for exploratory studies. Convergent validity is examined using the average variance extracted (AVE) and all constructs were above the benchmark of .5 (Fornell and Larcker, 1981)

With two modes and three tasks, it is possible to run this analysis 6 times. Figure 5 shows only one of them. We also ran this analysis with satisfaction as the dependent variable for the three tasks in search mode and the two remaining tasks in navigation mode. The results from this analysis can be summarized as follows. The following factors were significant: "got lost" every time, "level of difficulty" five of six times. "level of confidence" five of six times, negative emotions four of six times and positive emotions two times. Due to small size of the data-set we have to be cautious when interpreting these findings. Although it is positive that some of the factors seem to have a significant contribution each time or 5 of six times, the results must be interpreted as interesting and promising, but preliminary.

	AVE	Composite Reliability	Cronbach Alpha	Commun-ality
Got lost	0.82	0,93	0,91	0.82
Level of confidence	0,84	0,94	0,91	0,84
Level of difficulty	0,72	0,88	0,80	0,72
Negative emotion	0.74	0,92	0,89	0,74
Positive emotion	0,72	0,92	0, 88	0,72

Table 1: Constructs reliability assessment

5. Discussion and conclusion

In this explorative study, we set out to investigate users' preferences for two different retrieval strategies in large photo collections, search and navigation. A focus group was organized to gain insights about users' preferences, and a consecutive experiment with 20 participants was carried out.

The focus group showed that the participants, who were all experienced users and producers of digital photo collections, had clear preferences for navigation as the main retrieval method for their own personal photo collections. This is clearly in line with the recent literature in the field on search and navigation in the field of personal information management (e.g. Bergman et al 2008) where it is shown that users have a strong preference for navigation and with search considered to be "a last resort".

A closer look at the context and in particular the focus group participants' way of organizing their photo albums, reveal that their preference for navigation may be caused at least in part by the fact that their photo collections were not tagged in a way that could make search a realistic option. In other words, the photo albums were in practice not searchable. It seems that their preferences were shaped by this circumstance. However, taking the participants own explanations into account we can see that the contextual factors are only a part of a broader picture. The focus group participants themselves argued that navigation was preferred because this way of retrieving provided feelings of closeness, personal meaning and control that a search alternative would not provide in the same way. This is also in line with findings in the literature, e.g. Shneiderman (1997) who argues that navigation in hierarchies provides high degree of consistency as it confirms user expectations.

The participants in the experiment also appreciated navigation although they also meant search was the faster. The participants appreciated overview and feeling of control. One of the participants stated that: "It was easier to find a specific photo by searching, but I found it easier to get an overview with navigation. This is also what I am most used to in photo albums".

The initial results from the experiment showed that in this field explaining the role of perceptions and preferences is a complex undertaking. The Morae software, used in the experiment, maps the participants' behaviour, e.g time used on each task. We have seen that most of the participants in the experiments stated (when asked after they have completed the experiment) that they believed search was the fastest alternative. However, the measurement of time actually used on different type of tasks showed that navigation was the faster.

The analytical model of the experiment departed from the observation in recent research in the field of search and retrieval that there seems to be a lacking statistically significant relationships between various factors: mood prior to the search, topic of the search, sequence of task, task difficulty level, searcher's interest in a task and positive and negative affects. (Lopatovska 2009a) The present study addresses this challenge by combining two established and validated models, the subjective factor models used by Kules and Shneiderman (2008) and others, and the PANAS model that is applied in a number of recent studies of search (Lopatovska, 2009a and b). Both models are often applied separately, but to our knowledge not in an integrated way in studies of retrieval strategies. The rationale for this is that we believe applying both provide a richer picture of the role of subjective factors.

The results presented are based on a relatively small sample, but the findings indicate that there are significant relationships between the several of factors mentioned above and satisfaction with the task performance for various tasks and across different modes.

The results presented in the previous part showed that effects of emotions (moods) on satisfaction are less clear. Only partly significant relationships were found. Arguably, measuring the impact of emotions in this context may be challenging in several respects. The emotions reported may be shaped by a variety of factors in this experimental setting. Of course, interpreting the results in terms of "personal e-memory and retrieval" can be problematic since the participants don't have a personal stake in the outcome of the search and navigation tasks, as also observed and discussed by others ((Lopotovska 2009a). For student volunteers their main motivation for participating may be the reward, which they will receive anyhow. The participants reported feelings can in principle be attributed to other, undefined factors that are not related to the task. Thus, it is not a surprising result that there was a partly a lack of significant relations when it comes to the study of emotions as provided by the PANAS model. What could be expected, and also was found in the material, was that for example feeling of getting lost was significantly correlated with level of satisfaction. But this is understandable as the independent variable "feeling of getting lost" can much more easily be associated with the actual task performance. This underlines the needs for models that are more sensitive to tasks and contextual factors.

In our future work we plan to further develop revisions of the analytical model and apply it on a larger sample of participants. A larger sample would provide an opportunity for further analysis of the relations between subjective and objective factors. The results indicate that the main initial idea about the need for developing models that combined "subjective" and "objective" variables has been supported. Studies of preferences must be combined with studies of actual behaviour and context in order grasp the complexity of retrieval practices.

A factor that is not studied here, but which in our view can be fruitful and important to integrate in the model, is the strong *social* character of "photo management" and user behaviour. The important role of the social aspects, collaboration and sharing was a major finding in the focus group and appears to shape the way digital photos are managed and retrieved. These are aspects that should be integrated in the research framework.

6. ACKNOWLEDGMENTS

We thank our participants for their time and efforts. This work was supported by a grant from the Norwegian Research Council. Thanks to Wolfgang Leister and Trenton Schulz for valuable comments and suggestions.

REFERENCES

[1] Agosto, D. E. (2002). Bounded Rationality and Satisficing in Young People's Web-Based Decision Making. *Journal of the American Society for Information Science,53*(1), 16-27.

[2] Arapakis , I., Jose , J.M. & Gray, P.D. (2008). Affective feedback: An investigation into the role of emotions in the information seeking process. *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, *New York, USA*, 395-402

[3] Bergman, O., Beyth-Marom, R., Nachmias R., Gradowitch, N., Whittaker, S. (2008) *Improved search Engines and Navigation Preference in Personal Information Management.* ACM Transactions on Information Systems. Vol. 26, Issue 4.

[4] Chin W. (1998). The partial least squares approach to structural equation modeling. In GA Marcoulides (Ed.),

Modern methods for business research: 295-358: Mahwah

[5] Fornell, C. & D.F. Larcker, (1981) Evaluating structural equation models with unobservable variables and measurement error, Journal of Marketing Research 18 (1) (1981) 39–50

[6] Gwizdka, J. & Spence, I. (2006). What Can Searching Behavior Tell Us About the Difficulty of Information Tasks? A Study of Web Navigation. *Proceedings of the 69th Annual Meeting of the American Society for Information Science and Technology, Austin, TX,3*.

[7] Gwizdka, J. & Spence, I. (2007). Implicit Measures of Lostness and Success in Web Navigation. *Interacting with Computers*. *19*(3). 357-369.

[8] Herder, E. & Juvina, I. (2004). Discovery of Individual User Navigation Styles. In G. D. Magoulas & S.Y. Chen (Eds.), *Adaptive Hypermedia AH2004 Workshop on Individual Differences in Adaptive Hypermedia*. Eindhoven, The Netherlands.

[9] Kahneman, D. (2000). Experienced utility and objective happiness: A moment-based approach. In D. Kahneman & A. Tversky (Eds.) *Choices, Values, and Frames* (pp. 673-692). New York: Cambridge University Press..

[10] Kules, B., & Capra, R. (2008) Creating exploratory tasks for a faceted search interface. Paper presented at the Second Workshop on Human-Computer Interaction (HCIR 2008).

[11] Kules, B., & Shneiderman, B. (2008). Users can change their web search tactics: Design guidelines for categorized overviews. *Information Processing & Management*, 44(2), 463-484.

[12] Li, Y., & Belkin, N. J. (2008). A faceted approach to conceptualizing tasks in information seeking. *Information Processing & Management, 44*(6), 1822-1837.

[13] Lindlof, T. R., & Taylor, B. C. (2002). *Qualitative Communication Research Methods, 2nd Edition*. Thousand Oaks, CA: Sage

[14] Lopatovska, I. (2009a). *Does the mood matter?* To appear in the *Proceedings of the 2009 International Conference on Affective Computing and Intelligent Interaction.*

[15] Lopatovska, I. (2009b). *Searching for good mood: Examining relationships between search task and mood.* To appear in the *Proceedings Annual Meeting of the American Society for Information Science and Technology,* 2009.

[16] Lopatovska, I. & Mokros, H. (2007). Willingness to pay and experienced utility as measures of affective value in the information retrieval objects: Users' accounts. *Information Processing and Management, 44*(1), 92-104.

[17] Mackinnon, A., Jorm, A. F., Christensen, H., Korten, A. E., Jacomb, P.A., & Rodgers, B. (1999). A short form of the positive and negative affect schedule: evaluation of factorial validity and invariance across demographic variables in a community sample. *Personality and Individual Differences*, 27, 405-416.

[18] Marchionini, G. (1995). Information Seekers and Electronic Environments. In *Information Seeking in Electronic Environments* (pp. 11-26), Cambridge, MA: Cambridge University Press.

[19] McEneaney, J. E. (2001). Graphic and numerical methods to assess navigation in hypertext. *International Journal of Human Computer Studies, 55*, 761-786.

[20] Nahl, D., & Tenopir, C. (1996). Affective and cognitive searching behavior of novice end-users of a full-text database. *Journal of the American Society for Information Science, 47*(4), 276–286.

[21] NCC (2009) *Mariage project*. http//:marriage.nr.no

[22] Nahl, D. & Bilal, D. (eds.) (2007). Information and Emotion. *Medford, NJ: Information Today, I*

[23] Nunnally, J. C. (1978), *Psychometric Theory* (2nd ed.). McGraw-Hill, New York.

[24] Saracevic, T. (2007). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: Nature and manifestations of relevance. *Journal of the American Society for Information Science and Technology, 58*(3), 1915-1933.

[25] Shneiderman, B. (1997) *Designing the user interface. Strategies for effective human-computer interaction*. Addison-Wesley Longman Publishing Co. Inc.

[26] Tauscher, L. & Greenberg, S. (1997). How people revisit web pages: Empirical findings and implications for the design of history systems. *International Journal of Human Computer Studies, 47*, 97-137.

[27] Toms, E., O'Brien, H., Mackenzie, T., Jordan, C., Freund, L., Toze, S. (2008). Task Effects on Interactive Search: The Query Factor. *Workshop Pre-Proceedings In Initiative for the Evaluation of XML Retrieval (INEX) 2007*, 359-372.

[28] Wang, P., Hawk, W. B., & Tenopir, C. (2000). Users' interaction with World Wide Web resources: An exploratory study using a holistic approach. *Information Processing and Management, 36*, 229-251.

[29] Watson, D., Clark, L. A., & Tellegen, A. 1988. Development and validation of brief measures of positive and negative affect: The PANAS scale. *Journal of Personality and Social Psychology,* 54(6), 1063-1070.