Open access • Journal Article • DOI:10.1007/S00530-006-0064-7

Collaborative video searching on a tabletop — Source link

Alan F. Smeaton, Hyowon Lee, Colum Foley, Sinéad McGivney

Institutions: Dublin City University

Published on: 01 Mar 2007 - Multimedia Systems (Springer-Verlag)

Topics: User interface design, Multimedia search, Collaborative software and Interface (computing)

Collaborative Video Searching on a TableTop

Alan F. Smeaton, Hyowon Lee, Colum Foley, Sinéad McGivney

Centre for Digital Video Processing and Adaptive Information Cluster
Dublin City University, Glasnevin, Dublin 9, IRELAND

August 16, 2006

AbstractAlmost all system and application design for multimedia systems is based around a single user working in isolation to perform some task yet much of the work for which we use computers to help us, is based on working collaboratively with colleagues. Groupware systems do support user collaboration but typically this is supported through software and users still physically work in independently. Tabletop systems such as the DiamondTouch from MERL, are interface devices which support direct user collaboration on a tabletop. When a tabletop is used as the interface for a multimedia system, such as a video search system, then this kind of direct collaboration raises many questions for system design. In this paper we present a tabletop system for supporting a pair of users in a video search task and we evaluate the system not only in terms of search performance but also in terms of user-user interaction and how different user personalities within each a pair of searchers impacts search performance and user interaction. Incorporating the user into the system evaluation as we have done here reveals several interesting results and has important ramifications for the design of a multimedia search system.

1 Introduction

In almost all multimedia searching and searching in other systems, user considerations are based upon a single user performing a search, and working solo. Yet much of the work we do in our daily lives, including writing, editing, designing and information seeking, is collaborative in nature. The ubiquitous keyboard, mouse and screen interface which we all use means that any collaboration we do is either person-to-person or to a lesser extent is done in software rather than allowing us to collaborate directly on the task at hand. If we were able to use an interface which supported direct collaboration then design considerations for multiple users working collaboratively on an information seeking task would raise many interesting questions for system design.

In the work reported here we investigate the use of a DiamondTouch interface for direct collaborative searching through an archive of video footage. DiamondTouch [1] is a shared tabletop workspace which allows 2 or more users to interact directly and which physically supports direct user collaboration. Underlying our search system, called Físchlár-DT, is a sophisticated multimedia search engine which provides retrieval of video *shots* from a video library. The focus of the work reported here is an investigation into the interactions among collaborative searchers, working in pairs, who are collectively trying to search a video library for shots to match a shared information need. How do users interact with each other and how does system design influence this ? Does the level and type of interactions among users affect search performance and in what ways ? What kinds of user personalities are best suited to collaborative searching through a multimedia archive ? These are questions we set out to investigate as the answers will seriously impact the design and use of novel interfaces like the DiamondTouch, not only in information seeking, but in other tasks as well.

The work reported here builds upon prior work reported at the TRECVid evaluation in 2005. TRECVid is an annual forum for comparing the performance of video retrieval tasks and the retrieval performance of our system was reported earlier in [2]. For the present paper we have performed a detailed analysis of user interactions during the information seeking process through manual annotation of CCTV footage of all searches. We also performed detailed personality profiling and matching among our search pairs and we present these results and an analysis of how each affects others.

The rest of the paper is organised as follows. In the next section we present background material on collaborative searching, video searching and previous studies of group work and tabletop interfaces. In section 3 we describe the two versions of the Físchlár-DT system we have built and in section 4 we introduce the experiments we carried out to investigate issues which affect system design. Our results and analysis are presented in section 5 which is followed by a concluding section.

2 Collaborative Video Searching: Literature Review

The study in this paper is centred around the degree of focus on the end-user in each of the three main areas of collaborative information seeking, video retrieval, and groupware development. The subject of this paper comes from the intersection of the three areas which resulted in a novel tabletop system for collaborative video searching which we describe later. Therefore, although our work is based on these three established research areas, the novelty of our approach is the combination of these and the resulting Co-Located Collaborative Video IR system. Video retrieval is an area of research which focuses very much on the technological aspects with very little systematic concern for users; collaborative information searching and groupware development, on the other hand, deal with how to provide useful facilities for multiple users and thus usability and user concerns naturally receive more attention. In this section we review related studies in each of these three areas with reference to the degree of importance of user-centredness and in regard to the other two areas.

2.1 Collaborative Information Seeking: Moving from Individual to Collaborative Searching

The scope of concern of Information Retrieval (IR) has been greatly expanded by the realisation that we need to understand users and their search contexts as one aspect of wider information seeking [3]. Debate on the issue of relevance judgments, on which the core of IR is based, has increased the emphasis on user-interface issues as well as user-centred design and evaluation. While a single user, sitting in front of a computer searching for a piece of information has been the standard environment, the collaborative aspect of information seeking has also been addressed recently, shifting concerns onto systems and services that support multiple users' needs.

Hansen and Järvelin [4] have studied how people's information seeking activity is carried out in a patent office environment where information objects are used and passed to others, noting that collaboration among users happens frequently and is an important characteristic of information seeking. Similarly, collaborative processes that happen in a library setting were studied in [5] including asking a friend or colleague, consulting library staff for searching, and sharing a search outcome, with the conclusion that collaboration between users and between users and library staff is also a significant element of searching. These studies highlight the importance of supporting in some way the collaborative aspect of information seeking, asserting a move away from the traditional assumption that in most cases, information seeking is an individual process. Integrating collaboration further into a system's operation, Blackwell *et al.* [6] showed how a novel system allows its collaborators to discuss certain topics while moving around "statement tokens" (physical plastic objects with embedded radio-frequency ID tags) with topic labels. This activity generated continuous automatic query formulation with results updated to users, further guiding and promoting discussion. Hence query formulation was a by-product of collaboration (discussion and moving around the tokens) putting the collaboration activity at the centre of the activity and allowing the technical system to disappear into the background, generally considered as a phenomenon for good design [7]. We will see some more examples of user-interfaces in Section 2.3.

Effectively leveraging collaborative environments should use the *automatic recommendation* of items and using collaborative filtering, a system should keep a usage history of individual users and recommend other items that users with similar usage history have accessed. Collaborative filtering is a lively research area and has been successfully employed in different domains such as TV programme recommendation [8], movie recommendation [9], joke recommendation [10] as well as commercially successful book recommendations on Amazon.com. Collaborative filtering is a good example of how information seeking can be supported in a *collaborative* way.

The area of collaborative information seeking moves our attention from a conventional single-searcher paradigm (and all its subsequent design, implementation and evaluations based on that paradigm) to incorporating multiple users into searching. In our work, co-located collaborative searching on a tabletop, multiple users conduct a search task around a tabletop interface while talking, suggesting, discussing and interacting with each other. More general groupware issues as well as an overview of tabletops will be discussed in Section 2.3.

2.2 Video Retrieval: a Technology-Driven Approach to Content-Based Video Retrieval

The field of video retrieval is based on technologies which are still at an early stage of development and most of the research effort in the area concentrates on the underlying techniques for automatic analysis of video content. Understandably the user and issues of usability are not yet of major concern, although the shift of focus from functionality to usability seems to have started.

Studies on how automatic video analysis techniques could be usefully used by potential users and especially what kind of novel presentation/interaction tools in video abstraction would be beneficial, can be found in the literature. For example, various layouts for video keyframes for gist understanding and for searching have been studied in [11, 12, 13, 14, 15]. Selecting the right keyframes for presenting the result of a user's video query has also been studied [16] and interactive playback tools that allow efficient content-browsing [17, 18], interactive montages of map and timelines to visualise news video contents [19] have been evaluated. These experiments address some of the end-user's experiences with video retrieval systems, in other words, studying the constituent elements of a more complete, integrated, full multimedia retrieval system by attempting to explore and optimise the details of the best methods to use in these searching/browsing/presentation tools.

Incorporating some of these tools into more complete, integrated video search and browse systems are currently being investigated. A good example of this is the various interactive video search/browse systems developed in the TRECVid interactive search task [20] running annually since 2001. Within this task, systems have been developed which feature a wide variety of interface elements and although several are ingenious and novel and provide full features for the user to conduct a video search task, in all cases they focus on the technological aspects of video analysis for searching rather than on the users' task and usability needs. Genuine consideration of how these novel underlying techniques could be usefully used by end-user searching and browsing is still in its infancy and we are still a long way from stepping back and starting from a user's rather than the system's perspective. The technology-driven, feature-laden approach to the video retrieval field illustrates its immaturity [21], though some aspects such as shot boundary detection (SBD) and keyframe extraction are mature and stable. As research in this field matures, more and more sub-areas and component technologies in video retrieval will become stabilised and used for other, larger elements as a base technology as well as allowing a more usability and a user-centred approach to the area.

The Físchlár-DT system, the main subject of this paper, uses keyframes on a tabletop interface as the users' unit of searching and the system's unit of retrieval. While this is now common in the aforementioned systems, collaborative video searching on a tabletop where keyframes are free-floating and react to the collaborators' touch is new, and our focus is their collaborative benefit for search performance.

2.3 Groupware and TableTops: Emerging Technology and Multi-User Focus

Unlike video retrieval, the field of groupware has a reasonably long history of development. Often referred to as collaborative computing or computer-supported cooperative work (CSCW), groupware has been around long enough for user and usability issues to become the major focus of the field.

The benefits of systems designed for collaboration are well-known. For example a study by Stewart et al. [22] observed 72 elementary school children conducting a story-creation task using the KidPix application (designed for a single user) and the KidPad application (designed for collaboration) and they report an obvious collaborative benefit of increased communication among children when using KidPad, including soliciting help from partners, fewer verbal commands, and generally improved interaction.

Many important issues in designing such groupware were identified and discussed more than a decade ago [23], from organisational/social issues to user evaluation issues and the difficulty in designing a system using a designer's intuition alone. Concerns about workspace awareness for distributed groupware have also been studied in depth [24], guiding the focus to collaborators (users) rather than to the system. We will return to awareness issues for collocated collaboration on TableTop systems in Section 3.2.

Focusing on tabletop systems, several hardware and single display groupware platforms have been developed from simple reactive tables to experimental interaction with tables. Tabletops attempt to leverage the considerable experience people have with collaborating around tables in the home and work place. Typically a computer display is projected onto the tabletop surface from above (front projection) or below the tabletop (rear projection). Users' interactions are recorded using either a touch sensitive display or are recognised from a video camera placed overhead and using finger segmentation and tracking. These tables can afford a high level of workspace awareness and collaboration amongst users.

Examples of tabletop systems include the Scoop table [25] which displays sharable documents that belong to a user object such as a mobile phone when placed on the table, Habitat [26] in which a user becomes aware of another remote user's daily activities from a table that continuously displays representations from the remote user's table, and the Little Table [27] that projects distinctive LED lights attached to a physical object on the table. More sophisticated technology demonstrations of a tabletop include Dialog Table [28], Future Office Table [29], Office of the Future [30], Hitachi Tabletop Display [31], as well as Augmented Surfaces [32], ConnectTable [33] and InteracTable [34]. Amongst these, Dialog Table has in its design some specific tasks and techniques such as "Explore" (to display descriptive text, video or sound clips), "Relate" (to bring up related art works), and "Make" (to create an email or postcard); ConnectTable provides a table connection mechanism whereby attaching two tables logically merges the two table spaces and users can drag objects from one table to the other; and InteracTable also provides some built-in interaction techniques such as "passage" (any physical object that can be tagged and carry specific information which can later be used to retrieve on other tables), "take-and-put" (take one piece of information on one side of the table/wall, then locate it on the other side without having to continuously touching the surface) and "shuffle" (throw an information object and it lands on to the other person's side). There are also tables that use special tangible widgets including SenseTable [35] which uses a custom-made puck with dials to respond accordingly on a table, and the BlowAway collaborative game [36] in which users around a circular table blow air on a special input device that moves a virtual balloon on the table.

The applicability of several tasks on the tabletop have been investigated including games [37], planning [38] and supporting mealtime communication [39]. The authors in [37] developed a tabletop game in which users could control robots that move around the tabletop surface with digital features projected around the robot. The object of the game was for each robot to 'kill' the other and missiles and explosions were projected onto the tabletop surface according to the robots' position. A park planning system was developed in [38], in this application the task was to develop a layout for a public garden. Users could use either physical objects (i.e. models of trees etc.) or digital representations of these objects. The authors in [39] designed a system to support mealtime communication, through browsing of digital photos. These photos were projected onto specially designed plates which could be shared amongst users.

When multiple users interact with a single table, user-coordination becomes an important issue in tabletop design. This includes on which orientation text should be displayed on the table, which user should be allowed to interact with which objects, whether some areas on the table should be reserved for only a particular user, etc. Addressing this issue, LumisightTable [40] uses a directional view-controlling film called “Lumisty” to show different information on different orientations for different users around the table depending on their viewing angle. Single Display Privacyware [41] uses a shuttering sequence method that alternates different information on different frames on the table projection and users wear special glasses that filters only those frames on the table that are for them to see. While the above two tables take a hardware approach to coordinating usage on the table, software-based user coordination has been also considered. Noting that relying solely on social protocol is not sufficient in resolving usage conflicts on a tabletop, Morris et al. [42] suggested various software-based coordination methods such as “No Selection, No Touches, No Holding Documents” (a user can initiate a global change only if no user is selecing/touching/holding an object on the table), “Explicit” (initial document owner can grant other user to use it), and “duplicate” (a contested object duplicates itself so that multiple users can use it at the same time). User studies have identified properties of orientation [43], partitioning [44] and territoriality [45] that implicitly and explicitly help coordinate usage among users of a shared table. From these user studies, design guidelines were suggested for co-located tabletop development [46].

All of these tabletop systems are in their early stages of development compared to more established distributed groupware, but as can be seen they do focus on users and usability because the novelty of tabletop systems naturally brings in user collaboration as a central theme. However, no specific tasks for which such a table is to be developed have been proposed, no systematic user evaluation for tabletop use has been done, much less a consideration of the dynamics among users around a table or how individual users’ personalities influence overall collaboration and task performance.

At Mitsubishi Electric Research Lab (MERL), a tabletop called DiamondTouch [1] and its software toolkit DiamondSpin [47] have been used to develop a series of DiamondSpin table applications to experiment with tasks in different user scenarios. Figure 1 shows a typical DiamondTouch setup, where two users are sitting on receivers (embedded in their chairs) and the display is projected onto the tabletop using an overhead projector. By touching the tabletop with a finger, a circuit is completed which runs from the transmitter in the table through the user to the receiver and back to the transmitter, and thus each input on the table can be associated with a user.

UbiTable [48] explores the “semantics of personal and public space” by allowing its users to share the table space with their own laptop space, the software managing the private space (only viewable on his/her own laptop), the personal space (allocation of area on the table accessible only by a particular

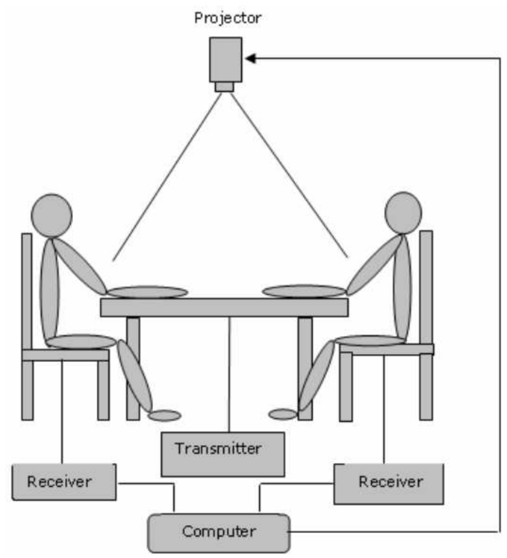

Figure 1: Typical DiamondTouch Setup

user but viewed by all users) and the public space (shared area with shared access by all users). The Opportunistic Browsing Coffee Table [47] displays a continuous stream of photos around the circumference of the table while the users browse photos in the middle of the table. Personal Digital Historian [49, 50] is a personal photo retrieval application that provides searching/browsing features based on who/what, where and when information. All of these applications use the DiamondSpin toolkit thus their documents and objects can be re-sized, rotated and moved around as free-floating on the table. The Físchlár-DT video search system which we have developed runs on a DiamondTouch tabletop, uses the same DiamondSpin toolkit and will be described in more detail in the next section.

3 Físchlár-DT: Searching for Video on a Tabletop

Físchlár-DT is a tabletop application we have developed running on DiamondTouch table [1] from the Mitsubish Electric Research Lab (MERL). The table is a horizontal, touch-sensitive retangular surface (107cm diagonal) with the chair seat as a receiver to detect which user is touching the table. Interactions from multiple users are simultaneously responded to with the identity of each user's interaction known. The table surface is robust enough to leave cups or note books on it while using the table. The table's output is via a high-resolution projector on the ceiling connected to the PC and vertically projecting the image downwards, filling the whole table surface.

DiamondSpin [47], a software toolkit accompanying the DiamondTouch, uses a real-time polar-to-

Cartesian conversion and thus virtual objects on the table can be rotated, re-sized and moved around. The Java-based toolkit provides objects for the table such as menu boxes, web browsers and pictures that can be rotated, re-sized and moved. The orientation schemes allow various location-sensitive automatic rotations whereby a virtual object on the table slowly turns its orientation from User A to User B as User A drags it from herself to User B. As DiamondSpin takes care of automatic orientations, an application developer can focus on task-oriented interaction on top of these provisions.

3.1 Video Retrieval for a Tabletop

Físchlár-DT is designed for two users (a pair) to collaboratively search for video shots from a video archive, currently containing 80 hours of broadcast TV news. The video archive is automatically indexed off-line, using a common shot boundary definition and keyframes provided by TRECVid. The system uses keyframes (representing a video shot) as the unit of retrieval, and the user can move these around and perform actions on them. Text generated from automatic speech recognition (ASR) is also indexed capturing the spoken dialogue during the video clip, and this is time-aligned with shots and keyframes allowing text searching to return associated keyframes. The keyframe search system uses the MPEG-7 Colour Descriptor and Edge Histogram Descriptor to calculate the similarity between two keyframes, and this finds all similar keyframes to a query keyframe when requested by the user. More technical details about the system can be found in [2].

3.2 Design for Awareness, Design for Efficiency: Interacting with Físchlár-DT

There are many issues related to Tabletop interaction that we could have investigated, such as what is the added benefit from allowing people to search together vs. searching independently. For the purpose of this paper however, we decided to focus on the collaborative interaction design of Físchlár-DT. In particular we assumed a scenario of two searchers working together and we focused on awareness of the workspace as it has been considered as one of the major issues in developing collaborative systems and its characteristics have been studied in [51, 24, 52]. Workspace awareness can be designed into software to varying degrees. Sometimes providing more awareness for a user actually results in less efficiency in the individual user's actions. For example, if we have a button to perform some function available for each user at a tabletop then use of the button near User A's side of the table can be easily overlooked by User B. If the action attached to the button causes a global change on the table, User B may be surprised at the sudden change; if the action causes effects only local to User A, User B might not be aware of the fact that User A has used it or how often User A uses it. Making the button action more easily noticeable by User B could be achieved by re-locating the button position to the middle of the table and providing only 1 button between the users, so that when User A uses it then User B knows. This simple re-design would result in better awareness between users, but User A needs to make more effort to reach out to the middle of the table thus reducing the efficiency of the individual interaction.

We have designed two versions of the Físchlár-DT application, one optimised for high *awareness* among users (at the cost of less individual interaction efficiency), and the other optimised for individual *efficiency* of interaction (at the cost of low awareness). Both versions of the application are used to search for as many video shots as possible within a given time, and we now describe how a pair of users interact with the each version of the system.

Two users sit on opposite side of the table, and one of them starts a search by typing in query terms into a text query box provided on the table, using a virtual keyboard. Users can move each of these elements to each other but only one set of query tools is provided. When a user taps the "SEARCH" button on the query box, the table is populated with 20 iconised keyframes whose spoken dialogue matched the text query terms. Using their fingers, users move keyframes around the table, or enlarge them for closer inspection by tapping and dragging the handle on the bottom right corner of a keyframe.

In the *awareness* version of Físchlár-DT, the table features 5 "hotspots" around the table that provide different functions when a user drags a keyframe over it:

- BROWSE: when a user drags a keyframe over this hotspot, the table is populated with another 20 keyframes from before and after the point in the video relative to the keyframe's location.
- PLAY: the video shot that keyframe represents is played on a separate external monitor.
- FIND SIMILAR: the table is populated with another 20 keyframes that are the most visually similar (using colour and edge histograms) to the placed keyframe.
- REMOVE: If a keyframe is not relevant or of use, a user can drag it to this hotspot and it will

disappear and not re-appear in future searches within the current task.

- SAVED AREA: a rectangular area on one side of the table is reserved as the ”saved area”. When a user finds a relevant keyframe, s/he drags it into this area and at the end of the session all keyframes in this area are saved as the search result from this pair’s search.

When a keyframe is dragged over a hotspot, a distinctive sound is heard making it obvious to both users what action has been triggered. Figure 2 (a) shows a snapshot of the table in the middle of searching. By having all major actions achieved by dragging keyframes to a particular location on the table and an accompanying characteristic sound for each action, awareness between the pair is highlighted. Although an equivalent highlighting of user awareness could be achieved by placing all hotspots in the middle of the table, the focus of the task here is on the browsing of keyframes and therefore these are displayed in the middle of the table. If we were to have placed keyframes on different parts of the table we may have seen users only examine keyframes closest to them.

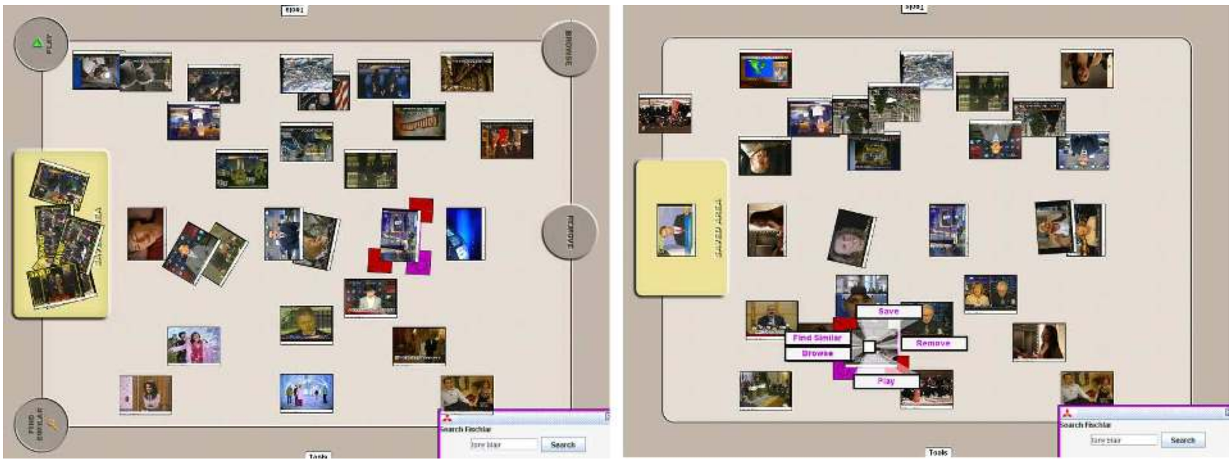

Figure 2: Tabletop for (a) *awareness* version and (b) *efficiency* version

In the *efficiency* version of Físchlár-DT, the same 5 functions are activated by a context menu which appears when a user double-taps on a keyframe. Figure 2(b) shows a snapshot of the table with a context menu shown. Note that there are no hotspots in this version, but the “saved area” is allocated a place on the table. When a user brings up the context menu and taps on “Save”, the keyframe will move to the saved area (i.e., disappear from the current location and re-appear inside the saved area). By using a context menu that appears just over a keyframe when requested and the response happens immediately, the user can work quickly and efficiently but with less awareness of the other user’s activities.

The purpose of developing these two versions of Físchlár-DT was to allow us to examine how these versions, with their highlighted and hidden levels of user awareness of each other, influence users' search performance and interaction.

4 Experiment

The "collaborative benefit", that is the benefit by having more than one user using any system together, is not a simple matter of having a greater amount of output from having more man-power used in the task. When co-operating during a collaborative task, individual users' personalities and social skills come into play in enhancing/reducing the collaboration thus influencing the performance of the pair. For simple tasks this benefit may be minimal, but importantly for tasks which require an increased amount of collaboration then bringing people to work together should provide greater benefit to the groups performance. We believe that search is one such example of a highly collaborative task. For our experiments, all users were chosen from within our research group and therefore were familiar with interactive video search applications, though none had seen or used the Físchlár-DT system before the experiment. An interesting future experiment might be to investigate how technical expertise influences group performance and therefore we may perform experiments with outside users with different levels of computer skills.

To address pair personality matching we used the Myers-Briggs Type Indicator (MBTI) which is sometimes used in studies to classify people into one of sixteen personality types. These use dimensions of

- Extraversion/Introversion,
- Sensing/Intuition,
- Thinking/Feeling,
- Judging/Perceiving.

Previous studies using MBTI have included studies of how personality impacts compatibility among pairs of student programmers [53], and in profiling the personality traits of security professionals [54]. The MBTI personality test was founded on Jung's theory of psychological type [55] and produces continuous preference scores for each of its four dimensions. These can be used to compute a degree of compatibility between individuals based on their scores in these 4 dimensions which can be calculated from a questionnaire. MBTI is not without its critics however, and many psychologists argue that MBTI is difficult to validate, that the terminology used is so vague that it allows any kind of behaviour to fit any personality type and that the attempt to pigeon-hole users into one of sixteen personality types is unfair. Although MBTI has flaws in terms of ability to measure personality and perform personality matching, for our purposes and because we are dealing with a small number (8) of pairs of users it was sufficient as an approximation of partial ranking of personality matching. Other personality measures would have been more exhaustive to compute — and regarded more highly — but were unlikely to differ hugely in the partial ranking of personality matching that we used.

A well-matched pair who communicate more efficiently, could co-operate in a time-saving manner and work more efficiently with each other and this should result in better performance and a better experience. Similarly, a poorly matched pair could result in poorer search performance and have a negative work experience due to their possible friction during their work, or possible dominance by one user. Yet just because two people have compatible personalities doesn't mean they will do better at searching for video shots. Two people who are both improvisers (perceiving) and both empathic (feeling) but whose personalities are both extraverted and idealistic may not be best matches for a long relationship, but might work well together in a searching task. The research question we address here is how does pair matching influence pair performance in the *awareness* vs. the *efficiency* versions of our system.

4.1 Experimental Environment - TRECVid

The TREC (Text REtrieval Conference) workshop series is sponsored by the National Institute of Standards and Technology (NIST) with the goal of encouraging research into the area of Information Retrieval (IR). Originally concentrating in the domain of text retrieval, in 2001 TREC introduced a video track (or task), in which particpants were required to develop systems to support content-based retrieval of video data. Just as in traditional (text) IR where the unit of retrieval was the text document, in video IR the unit of retrieval is the video shot. Since 2003 TRECVid has become a standalone workshop with empahsis on video IR tasks [20] . The goal of the interactive search task is for users to locate relevant video shots within the corpus of video data. Unlike other search tasks which concentrate on systematic improvements of the back-end search engine, the emphasis of the interactive task is firmly on the user and on developing intelligent user interfaces to enable users locate relevant material efficiently.

The 2005 TRECVid corpus of data consisted of 80 hours of broadcast TV News from various sources including CNN, NBC and Chinese and Arabic news networks. There were 24 multimedia topics, numbered 149 to 172, and for each topic there was a text description of the topic, some example images and some example video shots. Manual relevance judgments were made through a process known as ‘pooling’ in which results from each participating group were combined and assessors determined relevance for each of the results in the pool. These relevance judgments then became the ‘ground truth’ used to compute performance figures such as precision and recall for each of the groups, averaged over the set of topics.

4.2 Experimental Procedure

A total of 16 users (8 pairs) who had no knowledge of or experience with the system, participated in our experiments. Before the experiment, users were asked to complete an online Jung personality test to estimate their MBTI scores. The pairing of users was assigned prior to this.

Retrieval experiments were conducted in a room equipped with the DiamondTouch and a CCTV camera which recorded video and audio (see Figure 3 (a)). With an instructor present, each pair completed two training sessions followed by a training search task using each version of the system. Problems or questions on how to use the system were discussed during the training sessions. Each search task was 10 minutes long, and each task was to find as many shots as possible for one of the TRECVid topics. The 12 tasks were conducted by the pair, half of them with the *awareness* version, the other half with the *efficiency* version of the Físchlár-DT system, arranged using a Latin squares design to minimise bias and learning effect. The 12 tasks were spread over two separate dates so that users would not be overly tired when conducting the experiment and a post-experiment questionnaire was conducted at the end of the last search task. Figure 3(b) shows a photograph taken during one of the search sessions.

From the above procedure, the data we collected was (i) users’ personality type testing results; (ii) CCTV recording of all user interactions (including their faces and hands) which we annotated; (iii)

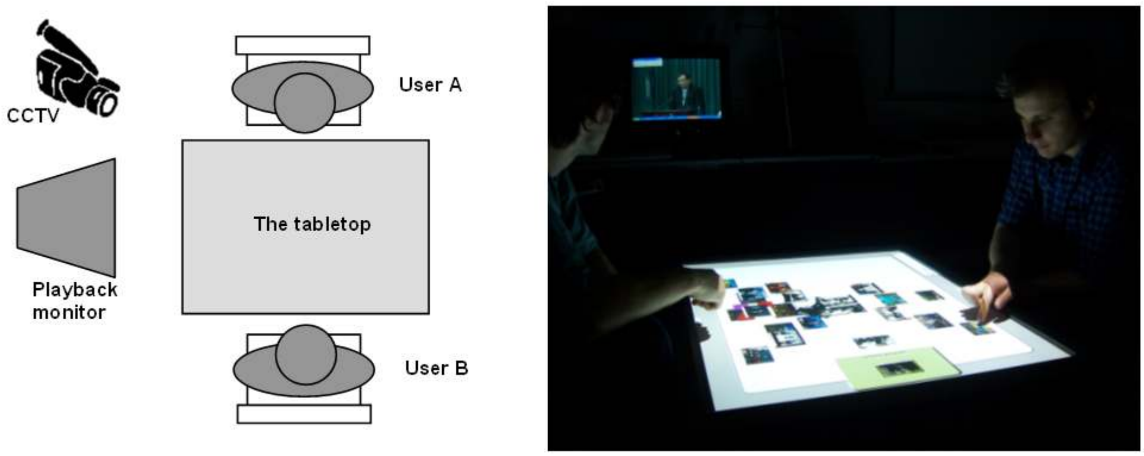

Figure 3: (a) Experiment Room Setting; (b) A Photo Taken During a Session

saved keyframes at the end of each search task, and (iv) user questionnaires on use of the systems.

4.3 Measures Used for Evaluation

We used the following measures to analyse the data we obtained in order to address our research question:

1. Pair match ranking - using personality MBTI scores to obtain a ranked list of best- to worst-match among the 8 pairs.

2. Amount and types of communications among pairs during search - we manually annotated the CCTV videos (18 hours total) that captured all user interactions with the table and among the pair using custom-built annotation software. This annotation time-stamped major events during search sessions in terms of the following, to be analysed and compared with pair matching and performance:

   - *Request/Response* – a *request* was noted if one user asks the other a question or requests him/her to perform an action; a *response* was noted when one user responded in some way to the other user's request. The modality of each of these were annotated individually as either voice, gesture or both.

   - *Comment* – If a user simply says something or makes a gesture that is not to elicit a response from the other user, it was annotated as a comment.

- *Coordination error* – this is an interaction error where either by mistake or by intention, one user hindered or disrupted the other user's search task. Examples are when a user triggered querying without notifying the other user, thus the sudden appearance of 20 keyframes surprised or bothered the other; when a user starts playback of a shot while the other user was already watching a shot; when a user enlarges a keyframe and it blocks other keyframes around that the other user was inspecting; when a user drags a keyframe that the other user was looking at, etc.

3. Search performance - we used the *treceval* software (a performance evaluation tool provided by TRECVid) to calculate Recall (percentage of relevant shots found by the pair) as well as Precision (percentage of saved shots that are relevant), to be used to calculate relative search performance among the pairs.

5 Results

Using the measures mentioned above, we explored the way these measures rank and correlate with each other. Pair matching against MBTI personality testing was used to rank how well each pair matched, and how this might be interpreted in terms of the actual interaction between the users during search sessions. We then examined pair match ranking against search performance and also correlated interaction among the pairs with performance and with pair matching.

5.1 Pair Matching

Using the results of the MBTI scores from each of our users, we ran compatibility testing to see how well the users in each pair matched. Table 1 shows the detail of each user's personality type and each pair's PersonMatch rating. For each user a percentage score is given, in parentheses, for each of the 4 personality dimensions namely Extroversion/Introversion (E/I), Sensing / Intuition (S/N), Thinking/Feeling (T/F) and Judging / Perceiving (J/P). We used an MBTI score generator to calculate the PersonMatch percentage. The pairs are ordered according to the PersonMatch percentage to produce our Pair Ranking. As we can see, only one pair matching (Pair A) was rated as "Good" (91%), one pair (Pair H) was "Unsatisfactory" (59%), and six other pairs were rated as "Satisfactory" (varying between 63 - 81%). There was no pair deemed as "Bad" (25-37%) according to the match testing. Our pairing strategy had been to pair people together who normally worked closely anyway and this probably resulted in no pairs that were ranked as "Bad " according to MBTI compatibility testing.

	Gender	Personality Traits	PersonMatch	Pair Ranking
Pair A	M M	E(82) N(45) T(60) J(44) E(11) S(12) T(12) P(11)	91% (Good)	1
Pair B	M M	E(33) N(50) T(12) J(78) E(22) N(62) T(38) J(38)	81% (Satisfactory)	= 2
Pair C	F F	E(67) S(1) F(25) J(33) I(22) S(25) F(62) J(44)	81% (Satisfactory)	= 2
Pair D	M M	E(44) N(50) T(1) J(33) I(22) N(62) T(62) J(56)	78% (Satisfactory)	4
Pair E	M M	E(11) N(62) T(12) J(33) I(78) N(62) T(75) P(11)	75% (Satisfactory)	5
Pair F	M M	E(11) N(38) F(75) J(56) E(67) N(62) T(38) J(67)	63% (Satisfactory)	= 6
Pair G	M M	E(56) N(31) T(1) J(22) I(44) N(12) T(75) P(11)	63% (Satisfactory)	= 6
Pair H	M M	I(100) S(38) T(25) P(11) E(11) N(50) T(12) J(33)	59% (Unsatisfactory)	8

Table 1: Personality Traits Types and the PersonMatch (How Well Each Pair Matches); 87-100%: Good, 63-86%: Satisfactory, 38-62%: Unsatisfactory, 25-37%: Bad.

5.2 Pair Interaction

As mentioned in Section 4.3, all interactions during search sessions were recorded on CCTV and manually annotated. Of a total of 96 sessions that the 8 pairs conducted, CCTV videos for 12 sessions were lost due to technical problems thus 84 sessions (amounting 14 hours of recorded video) were annotated and analysed. In the figures presented in this paper we averaged interaction figures over search sessions to account for missing recordings. Our expectation was that the amount and the different kinds of interactions that happened among users during search sessions would reveal or explain possible reasons for good/poor person matching and for good/poor search performance, in using *awareness* and *efficiency* versions of the Físchlár-DT system.

Figure 4 shows the average occurrence of interactions among the 8 pairs in all twelve 10-minute sessions, separated by *awareness* and *efficiency* versions. Different characteristics of communication among different pairs can be seen such as pair D showing the highest overall amount of communication for both systems among all pairs (average 61.5 times per search task for *awareness*, 70.5 times/task for *efficiency*), whereas pairs H, G and B show a relatively low amount (17.3 and 24.4 for Pair H; 26 and 28.8 for Pair G; 26.2 and 28.5 for Pair B). Pair E shows a much higher frequency of commenting (average 37.2 times/task in *awareness* and 26.5 times/task in *efficiency*) compared to their request-response rate.

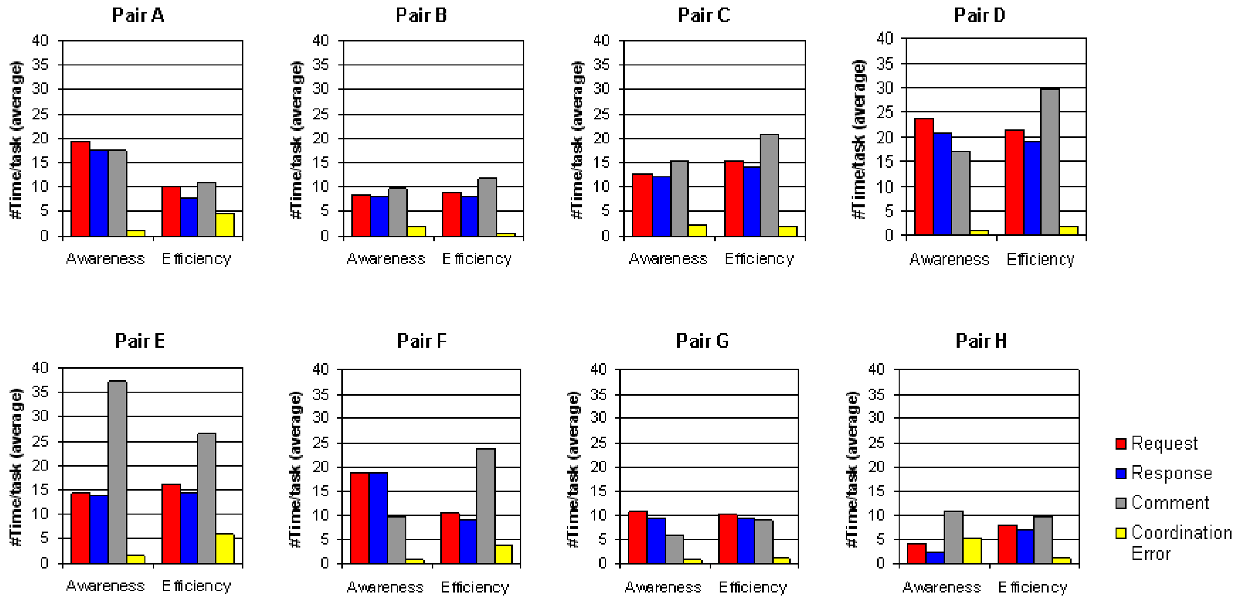

Figure 4: Average Amount of Pair Interaction and Coordination Errors by *awareness* and *efficiency* version per 10-minute search task.

There appears to be only a minor difference in terms of user interactions when using the *awareness* and *efficiency* systems across the eight pairs. This is illustrated in Table 2 where there is an almost identical number of overall interactions (1753 vs. 1734) with about 13% more requests/responses in the *awareness* than in the *efficiency* system.

As described in Section 3.2, in designing the two very different versions of the system in terms of provision of group awareness, we wanted to see in what way this difference influences search performances and pair satisfaction. One might expect the *awareness* version, by providing more cues of each user's actions and thus making collaboration between the pair more obvious, probably could reduce the need for explicit communication between users in conducting their searching. Pairs H, E and C seem to show that this is the case (Figure 4), but for other pairs this was not the case and for Pair D, F and A there were more communication between the users when using the *awareness*

	Awareness	Efficiency
Request	563	497
Response	520	441
Comment	608	702
Coordination Error	62	94
Total	1753 (per 42 sessions)	1734 (per 42 sessions)

Table 2: Total Amount of Interaction, Comment and Coordination Errors by System Type

version. One might also think the *awareness* version would result in less coordination errors between users because they are more aware of each other, but again this was the case only for some pairs (Pairs D, F, E and A) and there were pairs in which the *efficiency* version resulted in less errors (Pairs H, B and C).

To see if some of these phenomena could be better explained in the context of each user's personality traits, Figure 5 shows interaction characteristics by breakdown of each user for each system. Here we can see how User A and User B differ in their interactions with each other. For example, in the case of Pair D (which showed the highest level of communication among all pairs) User A requested more than he responded (requesting on average 12.7 times/task and responding 9.7 times/task), while User B responded more than he requested (average 9.7 vs. 10.3 times/task) on both versions of the system. This difference probably comes from the fact that User A is of an extrovert nature (+44%) and User B is more introverted (+22%) (see Table 1). The imbalance between request and response between the two users is in fact seen in all pairs to some degree. A pair's imbalance is even more emphasised in the *efficiency* version as Figure 5 shows. Pair C also shows the more emphasised imbalance between request and response in the *efficiency* version. On the other hand, for Pairs E and A the *awareness* version shows more degree of imbalance between request and response than in the *efficiency* version.

Pair E, who had the highest amount of commenting, shows both users commented more frequently when using the *awareness* version than the *efficiency* version, while User A (Extrovert 11%) commented more than User B (Introvert 78%) overall. Pairs D, F, G, B and C, on the other hand, show more comments overall when using the *efficiency* version than in the *awareness* version.

We now examine how the interaction between users within a pair changed over time, that is, as the pairs continue their tasks. As each person becomes more familiar with the search task, how to interact

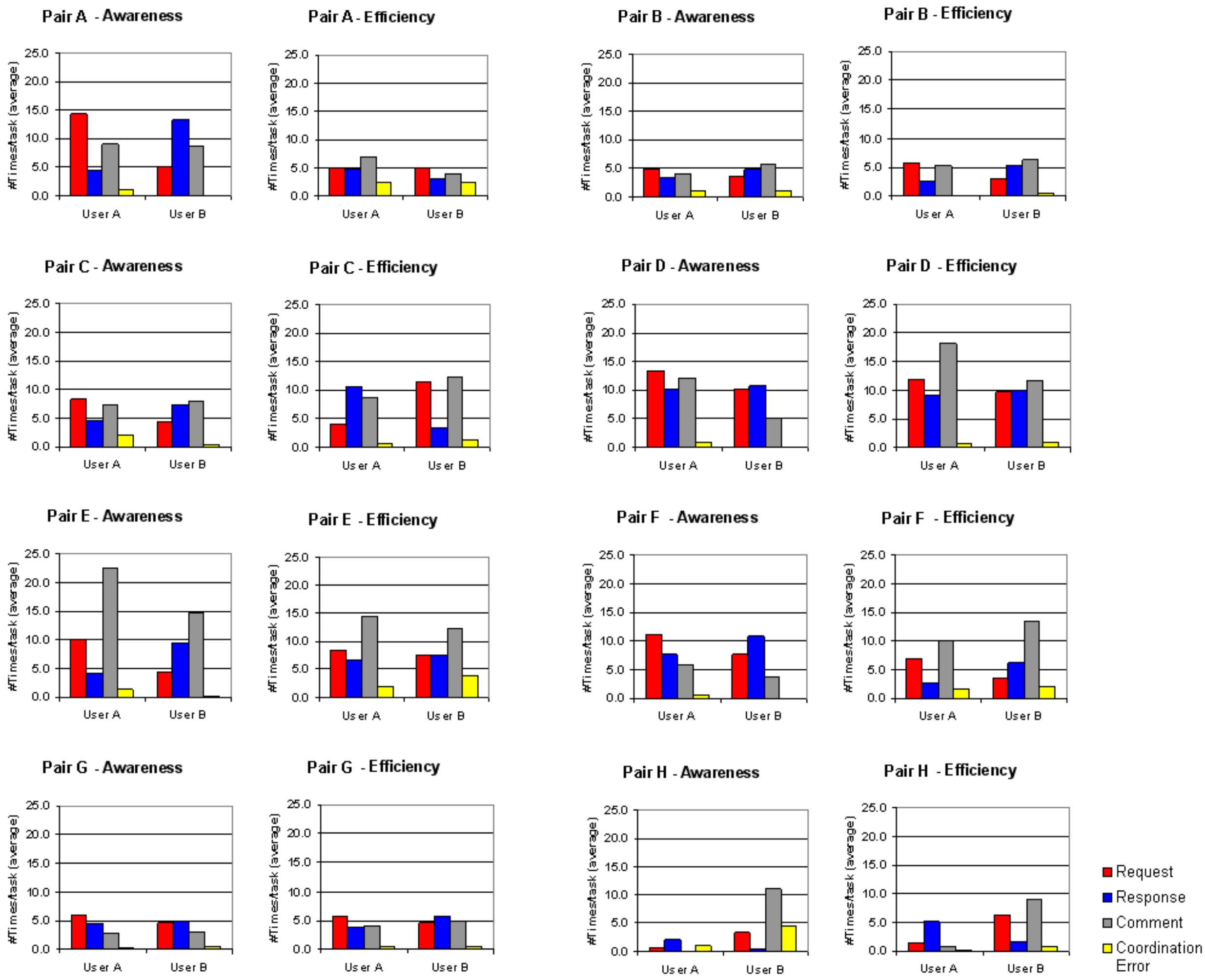

Figure 5: Amount of Pair Interactions and Coordination Errors by *awareness* and *efficiency* version (Split between Two Users)

with the system, the way the system responds to him/her, and indeed the person s/he is working with, s/he will be able to infer more and more about people's activities from more subtle cues [24]. One might expect the amount of explicit interaction with each other will become reduced as the users become familiar with each other's behaviour or habitual actions. We look at how this change could be differently manifested by the two different versions of the system in Figure 6.

The graphs in Figure 6 show the order of the 6 tasks each pair conducted using each of the two systems (x-axis), and the frequency of request/response (either voice, gesture, or both) comments and coordination error utterances. Due to some of the lost videos, pairs H, F and A show less than 6 tasks on the x-axis. For all cases, the occurrence of voice communication was much higher than gesture communication, with the voice request/response and comments having similar curves over time while gesture request/response was low without much change. For example, pair G shows that for the

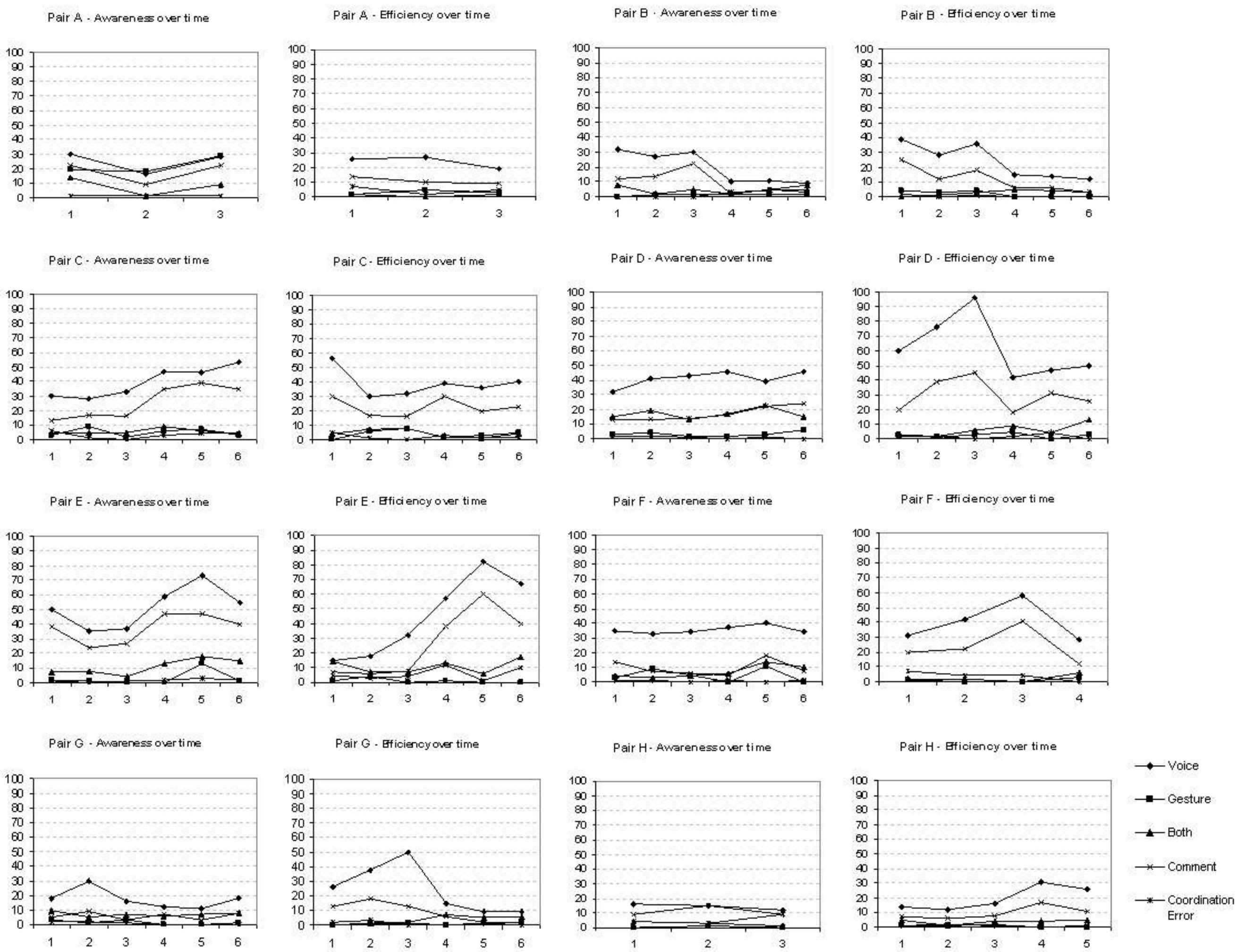

Figure 6: Amount of Pair Interaction; Request/Response (Voice, Gesture, or Both), Comment, and Coordination Error Over Time

awareness system, the pair gradually used less vocal requests/responses as the sessions progress while gestural requests/response were very small throughout. For the *efficiency* version, the reduction rate over time for voice is more prominent, with the last task ending up less than that with the *awareness* version. A similar phenomenon is observed for pair B.

Contradicting these, pairs H and E show increased voice communication for both systems as the sessions progress, while gesture was mostly low throughout except at the 5th session by pair E using the *awareness* version. As in Figures 3 and 4, there seems to be no instinctive pattern of interaction and coordination error between using *awareness* and *efficiency* versions by looking at them alone. We will relate these with pair matching (Section 5.1) and search performance (Section 5.3) in Section 5.4.

5.3 Search Performance

For our interactive experiments eight groups of users each completed 12 topics, 6 on the *awareness* system and six on the *efficiency* system. Table 3 shows the results of our interactive experiments. The performance of the two system are evaluated in terms of Mean Average Precision (MAP), Recall and P@10, i.e. precision after 10 retrieved shots.

As can be seen the *awareness* system outperforms the *efficiency* system on all measures. Due to the intensive style of TRECVid searching we had originally expected the *efficiency* system, with its symbolic interaction metaphor, to outperform the Awareness system, which required exaggerated user movements to activate system options. This was an unexpected result. We should also note here that the absolute performance figures for recall are low because of the large number of relevant shots for some topics, as noted in the next section.

Table 3: Interactive Experiment Results

Run Name	MAP	P@10	Recall
Físchlár-DT *awareness*	0.1529	0.7167	0.0685
Físchlár-DT *efficiency*	0.1372	0.6042	0.0673

Although the absolute performance figures of our results does not compare favourably with performance of the top-ranked other groups in TRECVid, relative performance of interactive search across participating groups in TRECVid is cannot be compared as among the participating groups' systems and experimental setting there are too many variables including system interface, number of users, users' technical skills in searching, users' familiarity with systems, duration of training and search session, and so on. The main reason for using performance figures in TRECVid interactive search is to assess within-group variations rather than cross-group, in our case Awareness Vs. Efficiency.

5.4 Correlations Among Evaluation Measures

For search performance, we used *recall* as the measure. Recall shows the number of relevant shots saved by the pairs from all relevant shots known in the database. In measuring pairs' search performances, we note that not all pairs conducted the same tasks as each pair was assigned 12 topics from the 24 in TRECVid according to a Latin squares experimental design. Some tasks are more difficult to find than others for the following reasons:

- Físchlár-DT's search facilities may favour particular types of search. For example, for finding a particular named person (e.g. "Find shots of Tony Blair", topic 153) the searcher will start by typing in the name of the person, and the search result will mostly be the keyframes of shots that mention the name in the dialogue; for finding something more generic (e.g. "Find shots of one or more palm trees", topic 166), typing "palm trees" in the query panel might not always retrieve keyframes showing palm trees because the dialogue may not actually mention them.
- Some topics have more relevant shots than other tasks. For example, topic 150 "Find shots of Iyad Allawi, the former prime minister of Iraq" had 13 relevant shots whereas topic 161 "Find shots of people with banners or signs" had 1245 relevant shots, making them easier to find.

In comparing search performances among pairs, we were able to directly compare only those pairs who searched the same topics. Table 4 shows this partial ranking of pairs by topics, along with average recall figures for each pair in parentheses.

To compare the performances of searching across the pairs, we logically combined the partial rankings from Table 4, i.e. D-H-E-F, C-G-A-B, D-H-B-G AND E-C-F-A, into one combined partial ranking giving us D-H-E-C-F-(G/B/A) where G/B/A represents a tied rank. We also used a simple weighting scheme in which each pair was assigned a weight by its rank (see last column in Table 4), then summed these weights to rank all pairs and this gave us a very similar ranking as the logical combination of partial orderings, with a Pearson correlation between the two of +0.818. The rank positions for each pair for these two rankings are shown in Table 5, along with our PersonMatch rank positions from Table 1. If we compare the rank positions for search performance using the logical combination of partial rankings against the PersonMatch ranking we get a Pearson correlation of -0.358. If we compare search performance rankings using the weighted combination of partial rankings against the PersonMatch ranking we get a Pearson correlation of -0.124. What this means is that in our experiment we found that there is actually a *negative* correlation between search performance, and how well our pairs are matched so in effect the more mis-matched our pairs are in terms of personality, the better they perform in searching.

Ranking	Topics 149/154	Topics 155/160	Topics 161-166	Topics 167-172	Weight
1	Pair D (0.308)	Pair C (0.128)	Pair D (0.082)	Pair E (0.069)	4
2	Pair H (0.240)	Pair G (0.108)	Pair H (0.045)	Pair C (0.068)	3
3	Pair E (0.209)	Pair A (0.094)	Pair B (0.033)	Pair F (0.063)	2
4	Pair F (0.185)	Pair B (0.093)	Pair G (0.023)	Pair A (0.055)	1

Table 4: Partial Ranking of Search Performance by Task and Pair

Pair	Ranking from Logical Comb.	Ranking from Weighted Comb.	PersonMatch Ranking
D	1	1	4
H	2	=3	8
F	5	=6	=6
E	3	=3	5
G	=6	5	=6
B	=6	=6	=2
C	4	2	=2
A	=6	=6	1

Table 5: Rank Positions for Search Performance and for PersonMatch

5.5 Analysis of Results

There are several interesting and unexpected results which we obtained in our experiments. Firstly, we found that the PersonMatch ranking of pairs does not correspond to their search performance. The pair that matched best according to PersonMatch (pair A) came up last in the rankings of search performance and the pair that was worst in PersonMatch (pair H) was second or third in terms of search performance. In fact there was actually a negative correlation between pair personality and matching, and search performance. That means that the better the personality match within a pair, the worse is the searching performance in our experiments. What this suggests is that the more mis-matched are pairs of searchers, in terms of their personalities, the more effective they are at searching, at least in our experiments. This could be because they had to work harder on their interaction.

The amount of interaction among pairs did not diminish as the number of searches they performed progressed. Taking into account that pairs had also completed some training topic searches, this indicates that they probably reached stable levels of person-person interaction, so performing more searches per pair would probably not reveal much change in the amount or types of interaction.

Table 2 showed that the *awareness* system had more coordination interaction than the *efficiency* version of the system but the *efficiency* version had more comments and coordination errors and this is as expected. What was not expected was that the *awareness* version yielded better search performance than the *efficiency* version. This version of the system was designed to make users aware of each other's activities, at the expense of efficiency of searching, and is strong evidence that collaboration among searchers can be beneficial to overall task performance. Analysis of the video recording of the search session suggests that users coordinated their actions more in the *awareness* version than the *efficiency* version. For example one pair decided on a strategy whereby one user on finding a potentially relevant shot would pass the keyframe to the other user and continue their searching while the other user would view the video shot on the monitor and decide whether to save the shot or not.

In terms of the measures we used, we note that the PersonMatch and the MBTI itself may not have been a reliable measure of personality or person compatibility for this study since it measures long-term compatibility between people rather than shot-term. On the other hand, the search performance measure (Recall) we used was the most appropriate measure of search effectiveness. As an exercise we conducted exactly the same analysis with an alternative performance measure of MAP (Mean Average Precision) and found out the resultant rankings are nearly same as when we used recall figures though there were some minor swapping of pair rankings, but nothing to correlate better with the PersonMatch rankings in any way.

Finally, upon examination of the post-experiment questionnaires completed by all participants, we found that the *awareness* system was most preferred by the majority of pairs with six out of the eight pairs preferring the *awareness* system, one pair preferring the *efficiency* version and the other pair conflicted in their preference. The reasons given for this majority preference were that the *awareness* system was easier to use, was more natural and intuitive, enforced coordination and collaboration more and made people more aware of the others' actions. This could be because there is novelty associated with being able to drag keyframes across the tabletop in order to place them at hotspots vs. the more traditional double-clicking though we don't have enough data to prove this. In some cases users felt that the *awareness* system was actually faster and more efficient to use which it was as we saw in Table 3, though this was not what we expected. Dragging keyframes was also a more natural action than using pop-up menu choices and had a more novel aspect making it more fun to use hence making it more preferable.

6 Conclusion

This paper has introduced a novel Co-Located Collaborative Video IR system, the basis of which is the intersection of three established research areas namely, collaborative information seeking, video retrieval and groupware development. The paper has presented a systematic user evaluation of a collaborative multimedia search tool, used by a pair of searchers to satisfy some information seeking goal. We have analysed use of the system in terms of overall system effectiveness (search efficiency), various types of user interactions, and the degree to which users' personalities impact each of these. Our results have revealed, surprisingly, that a system designed to make users aware of each other at the expense of individual search efficiency is not only preferred by users, but is actually more effective, suggesting that people work better when they work together. We also found that the degree to which people appeared to match in terms of personality compatibility appeared to be a reverse indicator of search performance and people who were less than ideal personality matches performed better when working together in our information seeking experiment. Although in terms of the scale of our experiment we consider this as far as any elaborate lab experiment could be, a greater number of experiments with a larger number of pairs may help support this result to a greater statistical significance. Better evidence for this should also come from a longer-term, longitudinal style study using a deployed, real-life application. Finally, our third major conclusion is that the amount of interaction people have with their searching "partner" does not seem to diminish as they perform more and more searches, although we have only been able to observe this for a total of 12 searches per pair. Although not shown in the results section, there is about a 10% increase in the amount of user-interaction as users move from their first search to their last, and this holds true for both the *efficiency* and *awareness* versions of our video search system. This suggests that users will continue to interact strongly with each other for both the social benefit, and the benefit of improving the performance in the information seeking task.

Each of these findings has ramifications for how we should design systems to support collaborative tasks in multimedia environments and collectively they show that supporting interactions among the users, in our case by making users aware of each other's task activities, has positive impact on overall task performance.

There are several avenues for future investigation into collaborative shared workspace multimedia searching and these include support for the division of labour of a task into sub-tasks and allocation of sub-tasks among users, as well as overall coordination of actions among users. Also in our plan for future work is developing a distributed version of Físchlár-DT where user(s) at more than one DiamondTouch in more than one physical location would collaborate in information seeking, with and without support from audio and video teleconferencing. In future work we intend to focus on each of these by designing different versions of multimedia search systems that emphasise these dimensions.

Acknowledgments

The authors gratefully thank the Mitsubishi Electric Research Lab (MERL) for use of a Diamond-Touch for these experiments. This work was supported by Science Foundation Ireland under grant 03/IN.3/I36.

References

[1] P. Dietz and D. Leigh. DiamondTouch: a Multi-user Touch Technology. In *UIST'01: Proceedings of the ACM Symposium on User Interface Software and Technology*, pages 219–226, 2001.

[2] C. Foley, C. Gurrin, G. Jones, H. Lee, S. Mc Givney, N. O'Connor, S. Sav, A.F. Smeaton, and P. Wilkins. TRECVid 2005 experiments at Dublin City University. In *TRECVid 2005: Proceedings of TRECVid 2005 Workshop*, 2005.

[3] G. Marchionini. *Information Seeking in Electronic Environments.* Cambridge University Press, 1995.

[4] P. Hansen and K. Järvelin. Collaborative Information Retrieval in an Information-Intensive Domain. *Information Processing and Management*, 41(5):1101–1119, September 2005.

[5] M. Twidale, D. Nichols, and C. Paice. Browsing is a collaborative process. *Information Processing and Management*, 33(6):761–783, 1997.

[6] A. Blackwell, M. Stringer, E. Toye, and J. Rode. Tangible Interface for Collaborative Information Retrieval. In *CHI 2004: Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 1473–1476, 2004.

[7] D. Norman. *The Design of Everyday Things.* Basic Books, 2002.

[8] D. O'Sullivan, B. Smyth, D. Wilson, K. McDonald, and A.F. Smeaton. Improving the Quality of the Personalized Electronic Program Guide. *User Modeling and User-Adapted Interaction*, 14(1):5–36, 2004.

[9] D. Cosley, S. Lam, I. Albert, J. Konstan, and J. Riedl. Is seeing believing? How Recommender Interfaces Affect Users' Opinions. In *CHI 2003: Proceedings of the Conference on Human Factors in Computing Systems*, pages 585–592, 2003.

[10] K. Goldberg, T. Robert, D. Gupta, and C. Perkins. Eigentaste: a Constant Time Collaborative Filtering Algorithm. *Information Retrieval Journal*, 4(2):133–151, 2001.

[11] A. Komlodi and G. Marchionini. Key Frame Preview Techniques for Video Browsing. In *DL'98: Proceedings of the 3rd ACM International Conference on Digital Libraries*, pages 118–125, 1998.

[12] T. Tse, S. Vegh, G. Marchionini, and B. Shneiderman. An Exploratory Study of Video Browsing User Interface Deisgn and Research Methodologies: Effectiveness in Information Seeking Tasks. In *ASIS'99: Proceedings of Annual Meetings of the American Society for Information Science*, pages 681–692, 1999.

[13] B. Wildemuth, G. Marchionini, T. Wilkens, M. Yang, G. Geisler, B. Fowler, A. Hughes, and X. Mu. Alternative Surrogates for Video Objects in a Digital Library: Users' Perspectives on Their Relative Usability. In *ECDL 2002: Proceedings of the 6th European Conference on Digital Libraries*, 2002.

[14] B. Wildemuth, G. Marchionini, M. Yang, G. Geisler, T. Wilkens, A. Hughes, and R. Gruss. How Fast Is Too Fast? Evaluating Fast Forward Surrogates for Digital Video. In *JCDL 2003: Proceedings of the 3rd ACM/IEEE-CS Joint Conference on Digital Libraries*, pages 221–230, 2003.

[15] M. Christel and A. Warmack. The Effect of Text in Storyboards for Video Navigation. In *ICASSP 2002: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2002.

[16] M. Christel, A. Warmack, A. Hauptmann, and S. Crosby. Adjustable Filmstrips and Skims as Abstractions for a Digital Video Library . In *ADL'99: Proceedings of IEEE Advances in Digital Libraries*, 1999.

[17] F. Li, A. Gupta, E. Sanocki, L. He, and Y. Rui. Browsing Digital Video. In *CHI 2000: Proceedings of the Conference on Human Factors in Computing Systems*, pages 169–176, 2000.

[18] W. Hurst, G. Gotz, and P. Jarvers. Advanced User Interfaces for Dynamic Video Browsing. In *MM'04: Proceedings of the 12th annual ACM International Conference on Multimedia*, pages 742–743, 2004.

[19] T.D. Ng, H. Wactlar, A. Hauptmann, and M. Christel. Collages as Dynamic Summaries of Mined Video Content for Intelligent Multimedia Knowledge Management. In *Proceedings of the AAAI Spring Symposium Series on Intelligent Multimedia Knowledge Management*, 2003.

[20] A.F. Smeaton, P. Over, and W. Kraaij. TRECVid: Evaluating the Effectiveness of Information Retrieval Tasks on Digital Video. In *ACM MM 2004: ACM International Conference on Multimedia 2004*, pages 652–655, 2004.

[21] D. Norman. *The Invisible Computer: Why Good Products Can Fail, the Personal Computer is so Complex, and Information Appliances are the Solution*, chapter 2. Growing up: moving from technology-centered to human-centered products, pages 23–50. MIT Press, 1999.

[22] J. Stewart, E. Raybourn, B. Bederson, and A. Druin. When Two Hands are Better than One: Enhancing Collaboration Using Single Display Groupware. In *CHI'98: Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 287–288, 1998.

[23] J. Grudin. Groupware and Social Dynamics: Eight Challenges for Developers. *Communications of the ACM*, 37(1):93–105, 1994.

[24] C. Gutwin and S. Greenberg. Workspace Awareness. In *CHI'97: Proceedings of the ACM Conference on Human Factors in Computing Systems, Workshop on Awareness in Collaborative Systems*, 1997.

[25] A. Hiltunen. Scoop: Information Reactive Table, Royal College of Art, London. Available at: http://myweb.tiscali.co.uk/anna.hiltunen/scoop/scoop.htm (last visited november 2005).

[26] D. Patel. Habitat: Awareness of Daily Routines and Rhythms Over a Distance Using Networked Furniture. In *LCS 2003: Proceedings of London Communications Symposium*, pages 145–148, 2003.

[27] M. Lecours. The Little Table, Concordia University, Available at: http://dfar_360.hybrid.concordia.ca/students/maxime/html/interaction.htm (last visited november 2005).

[28] M. Walczak, M. McAllister, J. Segen, and P. Kennard. Dialog Table. Kinecity Llc. Available at: http://dialogtable.com/ (last visited november 2005).

[29] Ars Electronica Futurelab. Future Office Table. Available at: http://www.aec.at/en/futurelab/future_office_table.asp (last visited November 2005).

[30] M. Billinghurst, R. Grasset, R. Green, and M. Haller. Inventing the Future Down Under: The Human Interface Technology Laboratory New Zealand. *SIGRAPH Newsletter*, 39, May 2005.

[31] Hitachi. Hitachi table top display. Aailable at: http://www.popgadget.net/2004/11/hitachi_table_t.html (last visited november 2005).

[32] J. Rekimoto and M. Saitoh. Augmented Surfaces: A Spatially Continuous Workspace for Hybrid Computing Environments. In *Proceedings of CHI'99*, 1999.

[33] P. Tandler, T. Prante, C. Muller-Tomfelde, N. Streitz, and R. Steinmetz. ConnecTables: Dynamic Coupling of Displays for the Flexible Creation of Shared Workspaces. In *UIST'01: Proceedings of the ACM Symposium on User Interface Software and Technology*, pages 11–20, 2001.

[34] C. Magerkurth and P. Tandler. Augmenting TableTop Design for Computer-Supported Cooperative Work. In *CSCW'02: Proceedings of the ACM conference on Computer Supported Cooperative Work, Workshop on Co-located Tabletop Collaboration*, 2002.

[35] J. Patten, H. Ishii, J. Hines, and G. Pangaro. Sensetable: a Wireless Object Tracking Platform for Tangible User Interfaces. In *CHI'01: Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 253–260, 2001.

[36] K. Cannon, N. Kavanagh, J. Lynch, T. Veale, and C Vipond. BlowAway: the Winds of Therslow. Available at: http://www.blowaway.org/ (last visited November 2005).

[37] M. Kojima, M. Sugimoto, A. Nakamura, M. Tomita, H. Nii, and M. Inami. Augmented coliseum: an augmented game environment with small vehicles. In *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, 2006.

[38] Y. Rogers, L. Youn-Kyung, and W.R. Hazlewood. Extending tabletops to support flexible collaborative interactions. In *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, 2006.

[39] K. Nishimoto, K. Amano, and M. Usuki. pHotOluck: a home-use table-ware to vitalise mealtime communications by projecting photos onto dishes. In *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, 2006.

[40] M. Matsushita, M. Iida, and T. Ohguro. Lumisight Table: a Face-to-Face Collaboration Support System that Optimizes Direction of Projected Information to Each Stakeholder. In *CSCW'04: Proceedings of the ACM conference on Computer Supported Cooperative Work*, pages 274–283, 2004.

[41] G. Shoemaker and K. Inkpen. Single Display Privacyware: Augmenting Public Displays with Private Information. In *CHI'01: Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 522–529, 2001.

[42] M. Morris, K. Ryall, C. Shen, C. Forlines, and F. Vernier. Beyond "Social Protocols": Multiuser Coordination Policies for Co-located Groupware. In *CSCW'04: Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pages 262–265, 2004.

[43] R. Kruger, S. Carpendale, S. Scott, and S. Greenberg. How People use Orientation on Tables: Comprehension, Coordination and Communication. In *Group'03: Proceedings of International Conference on Supporting Group Work*, pages 369–378, 2003.

[44] E. Tse, J. Histon, S. Scott, and S. Greenberg. Avoiding Interference: How People use Spatial Separation and Partitioning in SDG Workspaces. In *CSCW'04: Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pages 252–261, 2004.

[45] S. Scott, M. Sheelagh, M. Carpendale, and K. Inkpen. Territoriality in Collaborative Tabletop Workspaces. In *CSCW'04: Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pages 294–303, 2004.

[46] S. Scott, K. Grant, and R. Mandryk. System Guidelines for Co-located, Collaborative Work on a Tabletop Sisplay. In *ECSCW'03: Proceedings of Eurpoean Conference on Computer Supported Cooperative Work*, 2003.

[47] C. Shen, F. Vernier, C. Forlines, and M. Ringel. DiamondSpin: an Extensible Toolkit for Around-the-Table Interaction. In *CHI 2004: Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 167–174, 2004.

[48] C. Shen, K. Everitt, and K. Ryall. UbiTable: Impromptu Face-to-Face Collaboration on Horizontal Interactive Surfaces. In *UbiComp 2003: Proceedings of 5th International Conference on Ubiquitous Computing*, 2003.

[49] C. Shen, N. Lesh, F. Vernier, C. Forlines, and J. Frost. Sharing and Building Digital Group Histories. In *CSCW'02: Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pages 324–333, 2002.

[50] C. Shen, N. Lesh, and F. Vernier. Personal Digital Historian: Story Sharing around the Table. *Interactions*, 10:15–22, February 2003.

[51] C. Gutwin and S. Greenberg. Design for Individuals, Design for Groups: Tradeoffs Between Power and Workspace Awareness. In *CSCW'98: Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pages 207–216, 1998.

[52] J. Preece, Y. Rogers, and H. Sharp. *Interaction Design: Beyond Human-Computer Interaction*, chapter 4. Designing for Collaboration and Communication, pages 105–140. Wiley, 2002.

[53] N. Katira, L. Williams, E. Wiebe, C. Miller, S. Balik, and E. Gehringer. On Understanding Compatibility of Student Pair Programmers. In *SIGCSE '04: Proceedings of the 35th SIGCSE Technical Symposium on Computer Science Education*, pages 7–11, New York, NY, USA, 2004. ACM Press.

[54] C. Gates and T. Whalen. Profiling the defenders. In *NSPW '04: Proceedings of the 2004 Workshop on New Security Paradigms*, pages 107–114, New York, NY, USA, 2005. ACM Press.

[55] C.G. Jung. *Psychological Types. Collected Works of C. G. Jung, Vol. 6*. Routledge & Kegan Paul, 1971.