

Físchlár-DiamondTouch: Collaborative Video Searching on a Table

Alan F. Smeaton, Hyowon Lee, Colum Foley, Sinéad McGivney and Cathal Gurrin

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

ABSTRACT

In this paper we present the system we have developed for our participation in the annual TRECVID benchmarking activity, specifically the system we have developed, Físchlár-DT, for participation in the interactive search task of TRECVID 2005. Our back-end search engine uses a combination of a text search which operates over the automatic speech recognised text, and an image search which uses low-level image features matched against video keyframes. The two novel aspects of our work are the fact that we are evaluating collaborative, team-based search among groups of users working together, and that we are using a novel touch-sensitive tabletop interface and interaction device known as the DiamondTouch to support this collaborative search. The paper summarises the backend search systems as well as presenting the interface we have developed, in detail.

Keywords: Video Searching, interactive searching, tabletop interface

1. INTRODUCTION

The annual TRECVID exercise features collaborative benchmarking of a number of tasks related to video information retrieval using a shared dataset and a metrics-based evaluation. Of particular interest to us here is the interactive search task where systems which allow users to interactively formulate and re-formulate searches for video shots, are compared.

Over the last few years we have taken part in the interactive search task using an approach which allows a user's topic to be formulated as a combination of text query run against the automatic speech recognition output, and image query which is run against the set of keyframes for the video collection. In the context of the overall set of TRECVID participants this is a fairly common approach when using a standard interface combination of screen, keyboard and mouse. Our past efforts in TRECVID interactive searching have concentrated on modeling the user interaction and trying to capture this as best we can within the user interface.

In TRECVID 2005 we want to extend our previous work on modeling user searching and explore two new aspects of modeling the search process to see how such extensions can be accommodated while still using a standard back-end search system composed of text searching and image searching. In the first of these we want to explore an aspect of video searching not previously investigated, namely collaborative, team-based searching. In a real-world scenario where video searching is required, there are cases where a group of people can be asked to search for the same video clip(s), in parallel. Interaction among the group can happen through informal verbal communications, but there is no system support for such sharing. The second aspect which we are exploring in TRECVID 2005 is the use of a collaborative touch-sensitive tabletop display called the DiamondTouch and our work involves developing a set of interface interactions which support typical search interactions and testing these in practice. These two aspects of the user search process are clearly dependent and re-designing the interface is much more than just designing the interface widgets, rather it involves designing system support for the whole searching process including query formulation, refinement, marking and saving of relevant (and non-relevant) shots, and, in our case, system support for the interaction and sharing among collaborative users.

The remainder of this paper is organised as follows. In the next section we give an overview of the DiamondTouch in terms of hardware and user interaction. This is followed by a summary of the backend Físchlár system we have used in TRECVID 2004 and will use again in TRECVID 2005. In section 4 we outline the interface we have designed and built and in section 5 we outline the user experiments we are planning for our participation in TRECVID 2005. We conclude in section 6.

Author contact: E-mail: Alan.Smeaton@DCU.ie

2. THE DIAMONDTOUCH

2.1. DiamondTouch Hardware

Recent research into shared display groupware²¹ have resulted in a number of technology developments in multi-user input devices, such as SmartSkin,¹³ DViT²² and our platform, DiamondTouch¹¹).

DiamondTouch is a multi-user, touch-sensitive tabletop input device developed at Mitsubishi Electric Research Labs (MERL) as a research prototype. DiamondTouch determines touch location information on the table independently thus allowing each touch on the common surface to be associated with a particular collaborator. The table surface consists of an embedded array of rows and columns of antennas arranged in a diamond shape which act as sensors to identify the exact location of user touch inputs. There are currently two versions of the tabletop in operation, DT88 with a 79cm diagonal and DT107 with a 107cm diagonal, both operate with a 4:3 aspect ratio. The antenna pitch on both units is 5mm (both horizontally and vertically) and through a process of signal interpolation, an accuracy of 0.1mm can be achieved. Figure 1 below outlines a typical DiamondTouch setup. Each of the collaborators has their own receiver, usually connected to the user through a conductive part of the user's chair.

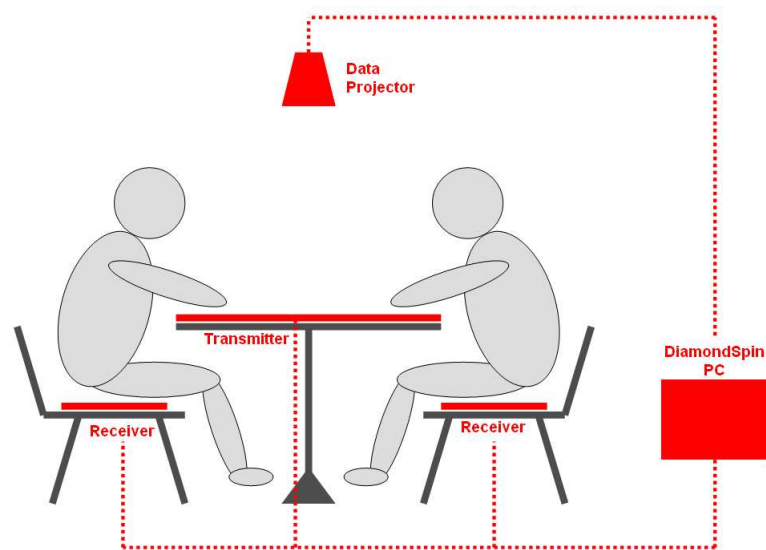


Figure 1. Typical DiamondTouch Setup.

When a user touches the surface a capacitively coupled circuit is completed. The circuit runs from the transmitter through the user to the receiver on the seat and finally back to the transmitter, thereby identifying the user. The DiamondTouch can support up to 4 users simultaneously without interference among their actions.

The tabletop nature of the display lends itself to a natural setting for collaboration; traditionally people meet and converse around tables. Interaction on the DiamondTouch is spontaneous and unencumbering rather than using additional devices such as a computer mouse or stylus, users are allowed to use their natural human tendencies of reaching, touching and grasping. The multilayer antenna arrangement allows collaborators to place physical objects on its surface while working without generating false touch points, whilst the robust and durable surface of the DiamondTouch table allows common coffee spills to simply be wiped clean.

2.2. DiamondSpin API

DiamondSpin¹⁹ is a software toolkit developed by MERL to support the rapid prototyping of multi-user collaborative systems. DiamondSpin has a well defined API of 30 Java classes and interfaces and uses pure JAVA 2D with JAI (Java Advanced Imaging) and JMF (Java Media Framework). DiamondSpin provides a real-time polar

to Cartesian transformation that enables arbitrary document positioning and orientation on a tabletop surface. The toolkit can be used to develop tabletop systems for rectangular, octagonal and circular tabletops. Unlike a traditional Cartesian coordinate system, a polar system has no predefined position for displaying documents (for example, keyframes in our case), in a polar coordinate system documents are displayed relative to one and only one meaningful centre. DiamondSpin uses a framework to describe every element in terms of its polar distance and polar orientation. This framework relies on two key concepts:

1. Translation of the origin of a conventional Cartesian display (usually the top left) to the centre of the table.
2. 3 Degrees of Freedom (DOF) d , α and β for each element on the display

The diagram in Figure 2 below shows a document being displayed at a distance d to the centre O at angle α , each document can then be rotated around its own centre using the β values.

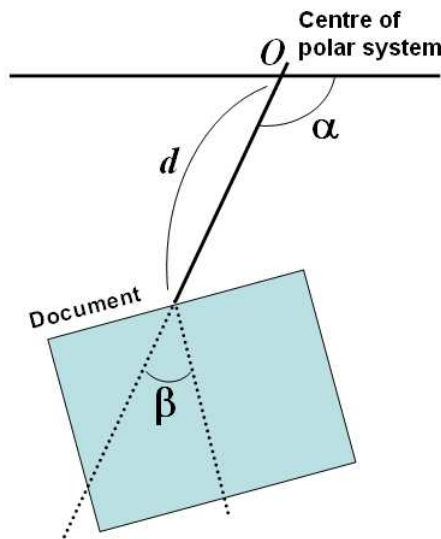


Figure 2. Document Rotation and Orientation.

The spontaneous nature of tabletop applications means there will be a high demand on the software to handle multi-user concurrent interaction. To enable this DiamondSpin uses a multi-layer representation with a multi-depth multi-thread repaint approach. Each DiamondSpin application uses 4 layers to display the elements of the system. The lowest layer, layer 3 contains non-interactive components such as background images. Layer 2 contains elements which may potentially become active but are currently not receiving inputs. Layer 1 contains active components, components which are currently receiving input events such as being dragged, resized or rotated. Layer 0 contains rotation sensitive components such as menu bars.

DiamondSpin uses a multi-thread approach to handle input events and repaint methods. One thread handles user interaction with tabletop elements, modifying element parameters such as position of the element, rotation angle of the table and the size of the element. Another independent thread handles repainting the UI, displaying changes to the surface.

The DiamondSpin API allows common Java factory components to operate on DiamondSpin's polar coordinate system, thereby allowing any UI component to be displayed at any orientation and rotated freely around the UI.

Eventually the toolkit allows the DiamondTouch application developers to use ready-made widget behaviours and common features for tabletop interaction and concentrate on application specific interaction techniques.

2.3. Related Work

Recent years have seen the emergence of a number of new hardware and software technologies aimed at supporting multi-user and shared display interactions. The *ConnecTable*²³ and *InteracTable*¹⁶ are part of the *i-Land* project being undertaken at the Fraunhofer Institute which uses rectangular tables that can be connected to each other thus expanding the workspace automatically. The emphasis here is on various office furniture being able to communicate with each other by seamless transition from their current proximity and usage. The issue of orientation and shuffling is only considered very briefly. *Escritoire*¹ from University of Cambridge uses 1 projector for focus (the document the user is working on) and another for peripheral information projected onto a large desk space, mainly to support a single user on a desk rather than supporting multi-users. *Lumisight*¹⁰ at NTT and University of Tokyo uses special projection hardware to display on a single table, different information to different collaborators by filtering the projection lights in different angles. *Single Display Privacyware*¹⁷ is a project which uses stereo glasses and different refresh rates of the display to display private information to each user, one problem with this approach to privacyware is encountered when a user attempts to point out a particular object in their private display to the other user. *DigitalDesk*²⁵ was a physical desk with camera based vision and projector capabilities thus merging the physical and electronic desks into one, the desk was designed for single users. Apart from the custom-built systems as the above, there are a number of *DiamondSpin*-based experimental applications built for *DiamondTouch* table. *UbiTable*¹⁸ is a project at MERL whose main emphasis is to develop and experiment the semantics of personal and public space. How 2-3 people gather at the table with their own laptop/PDA's and automatically connected to the table. Here the backend system manages personal and public information whilst sharing details of the collaborative interaction. *Personal Digital Historian*¹⁴ is concerned with the management of personal digital photos, searching is available over a users photo collection based on the 4W's (Who, Where, When, What). *Opportunistic Browsing Coffee Table*¹⁹ is another application using *DiamondTouch* and *DiamondSpin* here a continuous stream of photos are moving around the table, from which collaborators can opportunistically find interesting photos. *KidPad*²⁰ is a collaborative story authoring tool for children. In the project they developed 2 versions of the system, a single user system and a multi-user system. They observed dramatic differences when 72 local school children used 2 different systems.

3. THE BACK-END FÍSCLÁR-DIAMONDTOUCH SYSTEM

Físchlár-DT is based on the *Físchlár Digital Video System*, a platform on which we have been developing a number of video retrieval applications for different scenarios, including TV content browsing interfaces,⁷ educational video archiving for use by students,⁶ news update and archival in *Físchlár-News*,⁸ and a series of interactive search tasks for *TRECVID* experiments over the last 4 years^{43,5}. The system in 2005 has an XML-based architecture using MPEG-7 compliant video descriptions internally, and uses the same retrieval mechanisms as previous years' systems including text-based querying and relevance feedback using shot content.

The search test collection used in *TRECVID* in 2005 consists of 80 hours of broadcast TV news from CNN, NBC, MSNBC (English), LBC (Arabic), and CCTV4, NTDTV (Chinese), recorded during the last half of 1998. For the video content in a non-English language, machine translated English text was provided by NIST using off-the-shelf software. Along with the video data in MPEG-1 file format, we have made use of the common Shot Boundary reference from the Fraunhofer Institute, and the Automatic Speech Recognition (ASR) transcript provided by NIST.

3.1. Overall System Architecture

Figure 3 shows the architecture of the system, and will be explained throughout the following sub-sections. The back-end of the system is composed of offline indexing of the *TRECVID* 2005 video data, ASR text and its resultant MPEG-7 descriptors of various features (right half of Figure 3), and search engines for a text query and for keyframe similarity along with a result aggregation module (bottom left of Figure 3). The search server handles text and image based queries respectively and the rankings for text and image are aggregated to produce combined search result which is a list of relevant shots. The result is parameterised and passed to *DiamondSpin*'s search results display module, which deals with scattering the top 20 keyframes on the tabletop for the collaborators to browse.

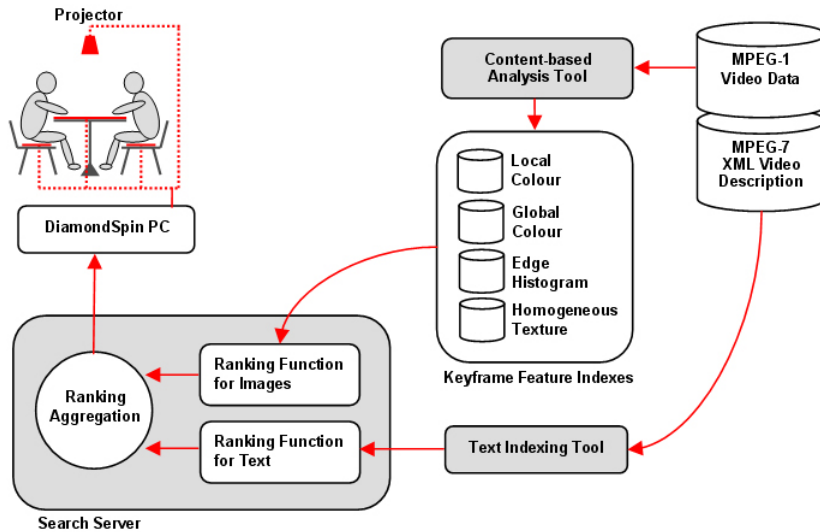


Figure 3. Físchlár-DiamondTouch System Architecture.

3.2. Retrieval and Weighting Scheme

There are two essential aspects of the retrieval and weighting scheme for the system: the text search and the image search. Both text and image ranking can be used singularly or in parallel. For example, for an initial query based on text only the text engine is queried, whereas for subsequent “Find Similar” queries (relevance feedback based on a keyframe), the combination of text and image engines are queried at the same time. We will discuss text and image search facilities, combination of feature evidences and the feedback mechanism in the following sections.

3.2.1. Text Search Facility

To support our text search facilities we used ASR text and in the case of non-English video, machine translation text. We implemented a search engine which uses the BM25 ranking algorithm with appropriate weights. The underlying text search engine (Físréal) was developed within the CDVP and used to support many other TREC experiments.² All text indexed was processed to remove stopwords and stemmed using Porter’s algorithm.

When processing a query, the system sends the query to the search engine to generate a ranked output (top 500) document list, which is then subjected to a smoothing phase where the scores of neighbouring documents were used to smooth the score of an individual ranked shot. The text search result can be used to rank shots before they are presented to the user (in the case of text based query), or alternatively combined with the results of an image search phase to generate a final ranked shot list (in the case of “Find Similar” relevance feedback query during the keyframe browsing stage).

3.2.2. Image Search Facility

In order to support keyframe matching we developed an image search engine which processed all keyframes using the automatically extracted feature descriptors described below. The four MPEG-7 descriptors were developed within the context of the aceToolbox, a toolbox of low-level audio-visual analysis tools developed as part of our participation in the EU aceMedia project (URL: <http://www.acemedia.org>).

Local Colour Descriptor (Colour Layout - CLD) is a compact and resolution-invariant representation of the colour in an image. The feature extraction process partitions an image into 64 (8x8) blocks and the representative colour of each block is determined. A DCT transform is then applied and a few low-frequency coefficients are selected using zigzag-scanning and nonlinearly quantized to form the CLD.

Global Colour Descriptor (Scalable Colour - SCD) measures colour distribution over an entire image. It is defined in the hue-saturation-value (HSV) colour space and produces a 256-bin colour histogram, normalised, non-linearly mapped into a four-bit integer value, and then encoded by a Haar transform.

Edge Histogram Descriptor (EHD) is designed to capture the spatial distribution of edges by dividing the image into 4x4 subimages (16 non-overlapping blocks) and then edges are categorized into 5 types (0, 45, 90, 135 and “nondirectional”) in each block. The output is a 5-bin histogram for each block, giving a total of $5 \times 16 = 80$ histogram bins.

Homogenous Texture Descriptor (HDT) describes directionality, coarseness, and regularity of patterns in images. It is computed by first filtering the image with a bank of orientation and scale sensitive (Gabor) filters, and then computing the mean and standard deviation of the filtered outputs in the frequency domain. In this work we only use the mean values to compute the similarity between the images.

More details on the four descriptors we used can be found in.⁹ The similarity between images was estimated by the L2 Minkowsky (Euclidean) distance for each of the features. The four features were the underlying methods by which we could compute keyframe to keyframe similarity.

At the time of image based query (“Find Similar”), the query image is used to identify similar shots by generating separate ranked lists of the top 500 shots for each of the 4 features. These ranked lists are combined into a single ranked list by examining shot rank position within the ranked lists as opposed to the actual similarity scores. This was done so as to avoid any combination issues due to very dissimilar score distributions among the features.

The overall outcome of the Image search element is a final ranked list of the top ranked 1,000 shots from all four features.

3.2.3. Relevance Feedback Mechanism

A central aspect of the Físchlár-DT system is that the user can select a keyframe to find other similar keyframes from within the database, by using relevance feedback. For example, if a user starts by querying for ‘ice hockey’ using a text query and locates a number of good examples of shots of ice hockey playing, then she can use each of these shots to further find similar shots in subsequent searches. Feedback works in different ways for text and for images. The ASR associated with the selected shot is used to query the image search engine and the four feature descriptors of the shot are used for the image search engine, and an aggregated ranking is produced. For text, any shot that is fed into the query contains candidate words that can be used to generate an extended query. We will feature a relevance feedback feature with multiple keyframes by allowing the collaborators to gather multiple keyframes and select them together as a group.

3.2.4. Combining Image and Text Evidence

When the “Find Similar” feedback feature is used, image and text evidence (ranked lists) need to be combined together to generate an overall ranked list for the system to present to the collaborators. This combination is done by rank position of shots from within both lists and the top 200 shots are returned for presentation to the collaborators. In situations where image is not required (e.g. text based query using the virtual keyboard) the ranked list was simply the top 200 shots from the text engine.

4. FÍSCHLÁR-DT INTERFACE

Each year our research group have taken part in the annual TRECVID benchmarking activity using some version of our Físchlár system for the TRECVID interactive search task. This year (2005) we will make use of the DiamondTouch hardware and DiamondSpin software to completely re-design the interface and interaction metaphor. This is completely different to our previous TRCVID systems, which saw a single user being given a specific topic to search for using a web-based front-end on a regular PC. The system we have developed encourages that the interaction with the system be both collaborative and much more physically natural by connecting the front-end to the DiamondTouch Tabletop instead of a desktop PC. Each user will be able to manipulate on-table objects

directly by using their fingers on the DiamondTouch Tabletop. Our motivation behind this is to make the task of searching for shots, given a pre-specified time limit, a collaborative one and to see how effective collaborative searching can be.

4.1. The Experimental Set-up

Our system allows for a two-person collaborative shot/image search as shown in Figure 4.

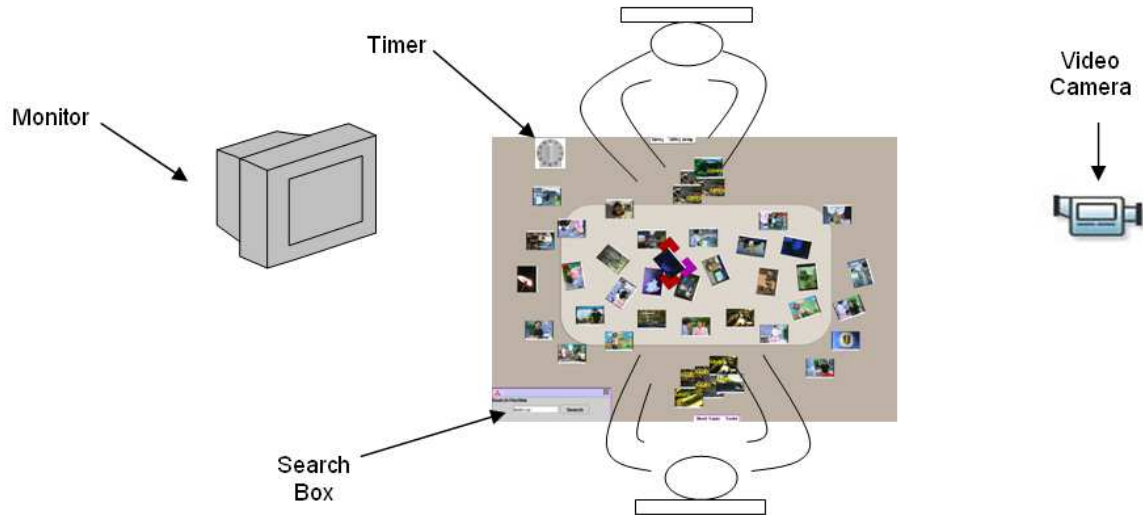


Figure 4. Typical Search Task Set-up.

As can be seen, users sit opposite each other and are supplied with a topic sheet to find as many shots relevant to the topic. A kitchen timer is placed on the Tabletop to enforce a time-limit on the search task, again reinforcing the idea of using real-world, natural physical objects. A video monitor is connected to the system to enable the video playback of shots selected by either user. In addition, we capture user interaction using video camera.

Previous TRECVID systems have facilitated initial text-based querying of Físchlár and this feature is retained through the inclusion of a text search box in the interface. This search box comprises a text-box and “Search” button to activate the search. However, only one of the users has control of this search box at any point in time. The text query is entered using a pop-up keyboard supplied by the DiamondSpin software toolkit. This is in keeping with the nature of the system - everything should be possible without the use of a peripheral device, just the use of fingers. Once a text query has been entered, the underlying search server returns a ranked list of images, twenty of which are scattered around the tabletop surface with the more relevant ones nearest the centre of the table.

For the tabletop we have decided to use a two-tone background to provide the illusion of each user having their own personal space. Simply by looking at the surface of the table, users can identify the central section (which is a lighter shade) as a kind of communal area where they can both manipulate images. The darker-coloured section however looks intuitively more like a private space for each user.

4.2. Image Manipulation

The DiamondSpin software handles the display and orientation of shots on the tabletop and allows them to be resized, rotated and moved. By tapping twice on a keyframe with a finger, a context-sensitive, pop-up menu appears (see Figure 5), enabling the user to search for similar images, save this keyframe, play the video shot back on the external monitor or remove this image from the tabletop.



Figure 5. Contextual Menu on a Keyframe.

The ability to playback a shot will require verbal communication and agreement between users to avoid any conflicts that may arise (i.e. both users wanting to play the shot at the same time).

Relevant keyframes/shots can be saved by clicking the “Save” option on the pop-up menu. A stamp stating “Saved” will then be placed on the top left-hand corner of the image. In order to find a similar shots to any given shot, he/she can then click on the keyframe and select the “Find Similar” option on the pop-up menu. This is a type of relevance feedback mechanism, and displays twenty more keyframes that look like this one as described in Section 3. Doing this will again require verbal agreement among users as only one “Find Similar” search can be executed at a time. The resulting images are then scattered around the Tabletop in a similar fashion to the initial text-based search results.

Figure 6 shows a table in the middle of searching, with some images saved and others still to be browsed.

As previously stated, the most interesting and novel aspect of our system is the collaboration of multiple users and how they communicate in order to avoid conflict. Much of this information will be captured by the video camera present in the room. By analysing video footage, we will determine the variation of the system that is most effective and most likely to provide sustainable results for a collaborative video search task.

5. EXPERIMENTS

At the time of writing we have developed and tested two variations of the Físchlár-DT described here and we have completed user experiments as part of TREVID. The two variations of the Físchlár-DT system both use exactly the same back-end search system, and the pool of 16 users we have used are all quite homogeneous, so



Figure 6. Table Screen in the middle of Search Task.

the variations are in the user interface and its support for collaboration, and in the resulting effectiveness as determined in TRECVID results, which are still being computed. The interface variations include the following:

- In designing a collaborative interface, multiple users using the same table resources can cause conflict with table space and use of elements on the table. For example, a user might want to view more closely a keyframe which is nearer to the other user, or a user might want to play a clip when the other user has just started playing another clip. While some observations revealed that the collaborators naturally divide regions on the table in order to do parallel work and thus conflicts seldom arise,²⁴ other observations state that frequent social mishaps appear in collaboration and thus solely relying on social protocols (“could I see your keyframe more closely?”) is not enough for smooth collaboration.¹² One of the interesting experimental settings we are considering is comparing two different styles of coordination policies: one in which all elements are open to all collaborators without any restrictions thus relying on social conventions for collaboration; another in which strict coordination rules are enforced on the table for example allocated shared and private areas on the table or once a user has started playing a clip then nobody else can request playing another clip until the first user stops.
- In providing collaborative video shot searching, we assume more than one user conducting in parallel the same search activity. However, there exist different ways of teamwork, such as working sequentially in tightly coupled activities, working independently, or working under assumed roles.¹⁵ Conventional search stages by a single user interaction can be divided into two, and each user allocated to different stages of action. For example, one user can start by querying and browsing to find initial candidate keyframes, and the second user can receive these keyframes from the first user and observe them more carefully (for

example playback each of the keyframes) to determine its relevance. There might be balancing problems in effort and time for each user, but this can be an interesting collaboration design that can be compared with working in parallel.

- Similarly to the above point, one way to maximise recall (the total number of saved shots) might be in parallelising the user's task: each user has his/her own text box and a separate search result appears for each user, with a shared area for saved shots. In this way, at a given time two users are searching separately thus (hopefully) doubling the output, at the cost of the level of collaboration between the users.

We have monitored the search sessions using video cameras and will annotate that footage looking for occurrences of evidence of these issues to correlate with overall efficiency of use.

6. CONCLUSION

This paper has described the Físchlár-DT system, developed as part of our participation in the TRECVID 2005 interactive search task. The system uses a conventional back-end search engine for text and image/keyframe searching, and a novel multi-user collaborative tabletop interface to support collaborative group searching. Our experiments, which have just completed, have explored the collaboration that exists, or must exist, among users working together to achieve retrieval of relevant video shots, and will relate this to overall system effectiveness.

ACKNOWLEDGMENTS

This work is partly supported by Science Foundation Ireland under grant 03/IN.3/I361, by the European Commission under contract FP6-001765 (aceMedia) and the Irish Research Council for Science Engineering and Technology.

REFERENCES

1. M. Ashdown, and P. Robinson, "Escritoire: a personal projected display," in *IEEE Multimedia*, **12**(1), pp. 34–42, 2005.
2. S. Blott, O. Boydell, F. Camous, P. Ferguson, G. Gaughan, C. Gurrin, N. Murphy, N. O'Connor, A.F. Smeaton, B. Smyth, and P. Wilkins, "Experiments in terabyte searching, genomic retrieval and novelty detection for TREC 2004," in *Proc. 13th Text REtrieval Conf. (TREC 2004)*, Gaithersburg, MD, 2004.
3. P. Browne, C. Czirik, G. Gaughan, C. Gurrin, G. Jones, H. Lee, S. Marlow, K. McDonald, N. Murphy, N. O'Connor, N. O'Hare, A.F. Smeaton, and J. Ye, "Dublin City University Video track experiments for TREC 2003," in *Proc. TRECVID 2003*, Gaithersburg, MD, 17-18 November, 2003.
4. E. Cooke, P. Ferguson, G. Gaughan, C. Gurrin, G. Jones, H. Le Borgne, H. Lee, S. Marlow, K. Mc Donald, M. McHugh, N. Murphy, N. O'Connor, N. O'Hare, S. Rothwell, A.F. Smeaton, and P. Wilkins, "TRECVID 2004 experiments in Dublin City University," in *Proc. TRECVID 2004*, Gaithersburg, MD, 15-16 November, 2004.
5. G. Gaughan, A.F. Smeaton, C. Gurrin, H. Lee, and K. Mc Donald, Design, "Implementation and testing of an interactive video retrieval system," in *Proc. 5th ACM SIGMM Intl. Workshop on Multimedia Information Retrieval (MIR2003)*, Berkeley, CA, 7 November, 2003.
6. C. Gurrin, P. Browne, A.F. Smeaton, H. Lee, K. Mc Donald, and P. MacNeela, "Físchlár-Nursing, using digital video libraries to teach processes to nursing students," in *Proc. IASTED Intl. Conf. on Web-Based Education (WBE 2004)*, Innsbruck, Austria, 16-18 February, 2004.
7. H. Lee, and A.F. Smeaton, "Designing the user-interface for the Físchlár Digital Video Library," in *J. Digital Information, Special Issue on Interactivity in Digital Libraries*, **2**(4), 2002.
8. H. Lee, A.F. Smeaton, and B. Smyth, "User evaluation outside the lab: the trial of Físchlár-News," in *Proc. 5th Intl. Conference on Conceptions of Library and Information Science -Context: Nature, Impact and Role (CoLIS5), Workshop on Evaluating User Studies in Information Access*, Glasgow, U.K., 4-8 June, 2005.
9. B.S. Manjunath J.R. Ohm, V. Vasudevan, and A. Yamada, "Color and texture description," in *IEEE trans. Circuits and Systems for Video Technology*, **11**(6), 2001.

10. 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 *Proc. ACM Conf. on Computer Supported Cooperative Work (CSCW '04)*, pp. 274–283, 2004.
11. Mitsubishi Electric Research Laboratories, DiamondTouch: A Multi-User Touch Technology, www.merl.com.
12. M. Morris, K. Ryall, C. Shen, C. Forlines, and F. Vernier, "Beyond "social protocols": multi-user coordination policies for co-located groupware," in *Proc. ACM Conf. Computer Supported Cooperative Work (CSCW '04)*, pp. 262–265, 2004.
13. J. Rekimoto, "SmartSkin: an infrastructure for freehand manipulation on interactive surfaces," in *Proc. ACM Conference on Human Factors in Computing Systems (CHI 2001)*, pp. 113–120, 2001.
14. C. Shen, N. Lesh, and F. Vernier, "Personal Digital Historian: story sharing around the table," in *Interactions*, pp. 15–22, 2003.
15. S. Scott, K. Grant, and R. Mandryk, "System guidelines for co-located, collaborative work on a tabletop display," in *Proc. European Conf. on Computer Supported Cooperative Work (ECSCW'03)*, 2003.
16. N. Streitz, J. Geisler, T. Holmer, S. Konomi, C. Muller-Tomfelde, W. Reischl, P. Rexroth, P. Seitz, and R. Steinmetz, "i-LAND: An interactive landscape for creativity and innovation," in *Proc. ACM Conf. on Human Factors in Computing Systems (CHI'99)*, pp. 120–127, 1999.
17. G. Shoemaker, and K. Inkpen, "Single display privacyware: augmenting public displays with private information," in *Proc. ACM Conf. on Human Factors in Computing Systems (CHI'01)*, pp. 522–529, 2001.
18. C. Shen, K. Everitt, and K. Ryall, "UbiTable: impromptu face-to-face collaboration on horizontal interactive surfaces," in *Proc. 5th Intl. Conf. on Ubiquitous Computing (UbiComp 2003)*, 2003.
19. C. Shen, F. Vernier, C. Forlines, and M. Ringel, "DiamondSpin: An extensible toolkit for around-the-table interaction," in *Proc. ACM Conf. on Human Factors in Computing Systems (CHI'04)*, pp. 167–174, 2004.
20. J. Stewart, E. Raybourn, B. Bederson, and A. Druin, "When two hands are better than one: enhancing collaboration using single display groupware," in *Proc. ACM Conf. on Human Factors in Computing Systems (CHI'98)*, pp. 287–288, 1998.
21. J. Stewart, B. Bederson, and A. Druin, "Single Display Groupware: A model for co-present collaboration," in *Proc. ACM Conf. on Human Factors in Computing Systems (CHI'99)*, pp. 286–293, 1999.
22. SMART Technologies, Digital Vision Touch Technology. White Paper, February 2003. http://www.smarttech.com/dvit/DViT_white_paper.pdf
23. 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 *Proc. ACM Symposium on User Interface Software and Technology (UIST'01)*, pp. 11–20, 2001.
24. E. Tse, J. Histon, S. Scott, and S. Greenberg, "Avoiding interference: how people use spatial separation and partitioning in SDG workspaces," in *Proc. ACM Conf. on Computer Supported Cooperative Work (CSCW '04)*, pp. 252–261, 2004.
25. P. Wellner, "Interacting with paper on the digital desk," in *ACM Comm*, **36**(7), pp. 86–96, 1993.