Mawd: A Collaborative Analysis Tool

Justin Ray Macfarlane

3 November 1995

For all my friends,
Both great and small.
Through all the bends,
I love you all!

Contents

1	Introduction	1
2	Overview	2
	2.1 HCI and CSCW	 2
	2.2 Ethnography	
	2.3 Design models	
	2.4 Observational Evaluation techniques	
3	Video's Role in Analysis	5
	3.1 Video's properties	 5
	3.2 Video's limitations	
4	Video Analysis Tools	7
5	A Collaborative Video Analysis Tool, Mawd	8
	5.1 Simplicity of use	 8
	5.1.1 Threads of Activity	
	5.1.2 Mawd's features	
	5.2 Quantitative results	
	5.2.1 Analysis of threads	
	5.3 Group awareness	
6	Discussion and Future Work	14
7	Conclusion	15
8	Acknowledgements	15
\mathbf{R}	eferences	16

Abstract

Video analysis is a powerful tool for studying human interactions. Transcribing video footage into usable and understandable data is tedious and time consuming. This report gives an overview of evaluation in HCI and CSCW, and describes the design and implementation of a Collaborative Video Analysis tool that eases data extraction and evaluation.

1 Introduction

The research fields of Human Computer Interaction (HCI) and Computer Supported Cooperative Work (CSCW) are concerned with improving how humans work with and around computers. Video analysis is a powerful tool for examining and understanding the interactions in HCI and CSCW. Video can capture all the subtle behaviours that occur in these interactions, with with negligible human effort. However there are severe problems in processing and analysing this data efficiently and in a balanced way.

This report describes Mawd, a tool that reduces the effort of video analysis, using threads of activity. Each thread represents a single type of action, such as Gaze, Talking, Typing and so on. Mawd logs user's actions and the resulting threads can be analysed. The design of Mawd has focused on three fundamental properties. First it must be simple to use. Second, it must produce Quantitative results. Third it must be group aware to efficiently support group analysis. This third property is unique: no other video analysis tool has addressed collaborative video analysis.

In this report, Section 2 gives an overview of evaluation and analysis. Section 3 details the characteristics video has when used for analysis. Section 4 gives a brief look at some other video analysis tools. Section 5 details the implementation and rationale behind the design of Mawd.

2 Overview

In this section, HCI and CSCW are introduced. Ethnography, an evaluation technique used in CSCW, is described. Finally Design models and observational evaluation techniques which are relevant to video analysis are discussed.

2.1 HCI and CSCW

HCI and CSCW are complex multi-disciplinary subjects that draw on many different areas, not just computer science. Techniques used are borrowed from Sociology, Psychology, Physiology and Linguistics to name a few.

HCI involves interaction at its heart and much of the work is involved in recording and analysing these interactions, to try to gain understanding of why they occur. HCI focuses on the individual's interactions, while CSCW looks at group interactions. Both seek improved methods of analysis and models of interaction. Both also rely on analysis to evaluate the effectiveness of the interface.

Analysis of recorded interactions is complex. Understanding human behaviour, documenting relevant characteristics and gleaning results is never easy. It is from interactions that Design Techniques, Models and Task Analysis stem, so the study of them in HCI and CSCW is of some importance.

HCI uses theories from cognitive psychology to gain understanding into interactions, CSCW has borrowed techniques from sociology (Hughes *et al.*, 1992). These techniques and theories are then adapted for the computer environment.

2.2 Ethnography

Ethnography is a sociological technique used to study groups (Rose, 1990) (Fetterman, 1989). It has become increasingly used in CSCW. In sociology ethnography's goal is to completely understanding and document how a group interacts, and why they interact as they do, through first-hand observation and analysis (Hammersley & Atkinson, 1983). Its basic techniques and ideas have been used in CSCW. The term "Ethnography" is used loosely in CSCW for anything that uses more than casual observation. Ethnography contrasts with the experimental psychology that is being used by HCI (Monk et al., 1993).

In CSCW's brand of ethnography tends for a broad understanding, and gaining as much information, for as little cost in time as possible (Shapiro, 1994). There are however in-depth studies of systems carried out (Suchman

& Wynn, 1984) and (Heath & Luff, 1982). Another difference is that sociological ethnography is almost completely a solo endeavour, while CSCW ethnographers work together in teams¹ (Bentley *et al.*, 1992). Often these teams involve both Computer Scientists and Sociologists (Harrison *et al.*, 1994).

Video is becoming an increasingly used method of observing groups in Ethnography. There is, however, surprisingly little literature explicitly on the use of video in ethnographical studies of CSCW systems.

Ethnography has been of particular interest in systems design (Hughes et al., 1994), where it is used for "requirement elicitation". It is used in two roles, as a part of the design process (Allen et al., 1993), and in the evaluation (Twidale et al., 1994).

2.3 Design models

HCI promotes iterative design as a method in the systems development lifecycle for interactive systems that produces a system that can more closely reflect the users needs. This is because the system is constantly being evaluated as to whether it fulfils the user's needs. This is generally done using prototyping. Successive prototypes are analysed, the results leading to the next prototype. Ethnography (and video) are becoming the method used in such designs.

Generally it is important to keep the user the centre of concern when designing interfaces, thus user-centred design (Nielsen, 1993), and make sure that their needs are met. Video is an ideal medium for representing users in such a manner (Brun-Cottan & Wall, 1995).

2.4 Observational Evaluation techniques

There are three methods of evaluating implementations, experimentally, observationally and with questionnaires (Dix et al., 1993). Of these observation is the least artificial. Observation is one of the most powerful methods of understanding user interactions, thus the popularity of Ethnography. There are various techniques that are widely used in HCI to evaluate how the system is being used:

• Think aloud is a good method of understanding how the user uses the system.

¹A large proportion of sociological ethnography books are solo efforts, while the CSCW papers are almost always done by a team

- Protocol analysis involves recording and studying the user as they use the system.
- Cognitive Walk-throughs can give a clearer insight than think aloud techniques, as it is post-task. It is also easier for the user during the task.

Ethnography and video analysis focuses on protocol analysis, as this is observing the user in their environment.

3 Video's Role in Analysis

Video has many properties that make it ideal for analysis. It does however has several severe limitations.

3.1 Video's properties

Video is the only way to obtain a thorough recording of events, that allows indepth observation (Roske-Hofstrand, 1989). It can provide a complete record of events (Carter & Anderson, 1989) that other methods of observation, such as note-taking, can not provide (Neal, 1989).

Video is a rich-medium which captures all aspects of an interaction that are within the frame. It offers a visual and audio record of events. There is a large depth of detail captured, which can be manipulated so that all desirable aspects can be observed.

The most valuable aspect of video is that it is reviewable. This is useful for many reasons. While a note-taker of an interaction can get quite a good insight into what is occurring, if their focus is in the wrong place then they may miss something. Being reviewable, video need not miss anything. Comparing what can be observed with notes, and video, it is clear that a lot will be missed (Allen, 1989). If a specific action is found to be of interest it can be viewed again, to gain all the detail necessary. Video provides a permanent record of the interaction.

If new theories on an interaction are developed, they can be tested on past video material, rather than having to make further observations (Neal, 1989). This saves a lot of effort.

Although video is contextually deep, it can be viewed with varying levels of detail. It is sensible to have several viewings to gain an over-view of what is happening, before more detailed viewing for analysis.

Video is inherently unbiased. When Note-taking, the observer will have an idea on what they are looking for before-hand, so can be biased. Things are overlooked, since they are not expected. Even if things are missed in initial watchings of video, there is always the opportunity to review the tape to catch them later.

Video is ideal for capturing group and inter-personnel interactions (Horton et al., 1989). It can capture the nuances of how people interact together, much better than someone observing can, after analysis. It is widely used in collaborative system evaluation (Tang, 1991) (Horton et al., 1989)

Frank understanding of users reactions to a system can be gained through video. Questionnaires can often be misleading to an analyst, because the answers given can reflect what they perceive as being the wanted answer,

rather than what they actually feel (Horton et al., 1989). Video can capture the body-language of the irritated, or confused much more obviously.

Video allows more interaction in the analysis. The fieldwork done in studying tends to be solitary. An analyst will go into the field and make their observations. This requires specialist skills and techniques. It is also not usually practical to have a team of analysts in the field, as this would be inconvenient to the subjects, and possible intimidating. Video analysis can be done with no-one present, while the tapes can be viewed by a group. This allows a larger range of views on what is a subjective task, adding balance. Reviewing a tape with others is a good way to test what has been observed (Allen, 1989).

3.2 Video's limitations

Video does have limitations however. These stem from the amount of detail possible. Analysis must be done bearing in mind the depth of analysis necessary. It could be easy to go into unnecessary and time-wasting depth.

Transcribing videotape can be tedious (Neal, 1989) and is time-consuming. Allen (1989) states it takes two to ten hours to transcribe each hour of videotape. It is very easy to capture hours of data, but even minutes of tape requires hours of analysis.

Another problem is in deciding what needs to be recorded. Enough data needs to be collected, but if too much is recorded, then analysis will prove costly (Allen, 1989).

The results of analysis are almost always qualitative. It is often useful however to gain a quantitative measure of some events. Qualitative results are harder to understand and gain meaning from. Quantitative results can be statistically analysed, thus giving a measure for comparison. Gaining quantitative results from analysis is not easy, as the methods of analysis currently tend to lead to qualitative results, and the interactions themselves are not easily quantified.

4 Video Analysis Tools

There are various tools that have been designed or proposed to complement video analysis. These tools use different strategies to aid in video analysis. These are all variations on a note-taking scheme. The following are four that have been more successful.

- EVA (Experimental Video Annotator) (Mackay, 1989) uses a multimedia workstation to annotate the video. This is done as the video is recorded. Interesting events are logged, snapshots can be taken, or time-stamps used. The act of note-taking however can distract attention off the events themselves, causing them to be missed. Its advantage is that it simplifies analysis considerably.
- The Workplace project (Trigg, 1989) This tool uses a range of different data streams. This means that is does not presuppose a single analytic framework. It is designed to incorporate whatever relevant techniques are required. It is time-line based, but loosely, and it is possible to zoom in and out, to gain finer or broader detail. It is meant to be flexible. Using this tool however is time-consuming, with all the various different measures being used.
- Marquee is a pen-based video logging tool (Weber & Poon, 1994). It allows the analyst to correlate their notes with the video. It is also possible to use it to take real-time notes (hand-written into the system).
- Video Transcriptor is a tool that allows transcription of digital QuickTime video (Weir, 1994). It allows the annotating of events that tie directly into the video.

5 A Collaborative Video Analysis Tool, Mawd

Mawd is designed to address three readily identifiable problems with video analysis, and its tools:

- 1. Non-trivial to use. Video-analysis tools are designed for expertusers, by expert-users, and have complex functionality.
- 2. Qualitative results. Ethnographic analysis produces qualitative results, which are much harder to analyse than quantitative.
- 3. Single-user design. All current video analysis tools are single user.

This section details how Mawd overcomes these problems. Subsection 5.1 details its simplicity of use. Subsection 5.2 deals with Mawd's quantitative results. Subsection 5.3 outlines the collaborative nature of the system.

In this section, threads of activity are discussed. The features of Mawd explained, including its group awareness and the possible uses of threads in analysis.

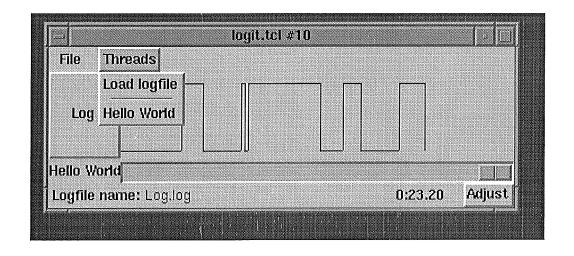


Figure 1: Mawd

5.1 Simplicity of use

The use of threads and a simple interface make Mawd a simple tool to use.

5.1.1 Threads of Activity

Ishii et al's use of threads in analysis of gaze awareness motivated this research (Ishii et al., 1992). In their analysis they used a graph of threads² to illustrate gaze, voice and hand gestures (See Figure 2). These threads where painstakingly compiled by hand. The results of analysis of these graphs (again by hand) was used to show the effectiveness of their successive prototypes in facilitating more eye-contact.

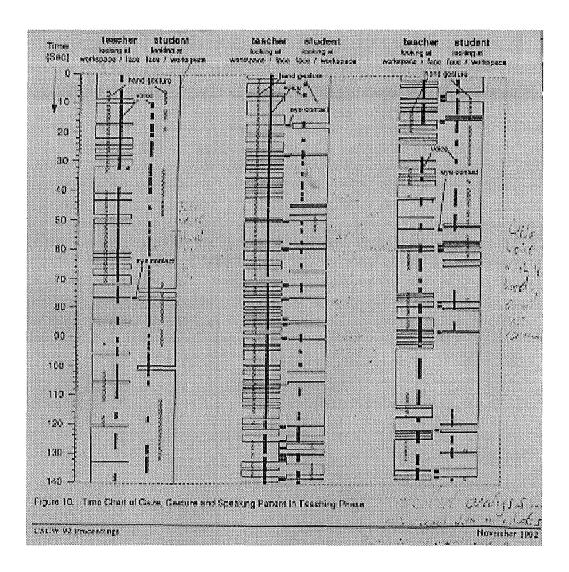


Figure 2: The threads used by Ishii et al to analyse gaze, gesture and talking.

²They described it as a time-chart

Threads are binary graphs that indicate activity or inactivity in relation to the simple action they are describing. They are an easily comprehendable measure of regularly occurring events.

Threads give Mawd its quantitative nature and make the tool simple in concept.

5.1.2 Mawd's features

Mawd is designed to be used while watching the video. As the video plays, the analyst logs an action under a thread. When the action occurs, they press the log button, causing a change in state for the thread. The thread then reverts back to the inactive state when the button is released. This can be observed by the user, as the thread is updated graphically in real-time.

The time-scale is manipulated by the slider, or under a menu. The slider can be used to scan back and forward through the thread. The menu moves the thread to that position. The use of a slider and time adjust is not consistent with the video metaphor however (Rieman et al., 1994). To keep the metaphor of a video, it would be better to use fast-forward and rewind buttons. The computer's ability to instantly move through the thread is taken advantage of, rather than the linear scrolling of videotape.

Another feature that is essential is the ability to edit threads. Video is reviewable, so the threads created should also be easily edited. Figure 3 shows a thread in the process of being edited. The original thread is in the top window. The time is adjusted to the required place and then the thread is logged over. The edited thread appears as a bolder line, as illustrated in the middle window. The old thread is not replaced, until the view is refreshed, as in the bottom window.

5.2 Quantitative results

The threads are a quantitative measure of actions on the video. This allows quantitative analysis of them.

5.2.1 Analysis of threads

Threads of course need to be analysed. Apart from simple measures such as the percentage of time doing the action, it is possible to combine threads.

If two threads are combined, then the intersection gives overlapping behaviour. The best example of this being of interest is the threads are recording gaze. Here the intersection indicates that eye-contact is occurring (Ishii & Kobayashi, 1992).

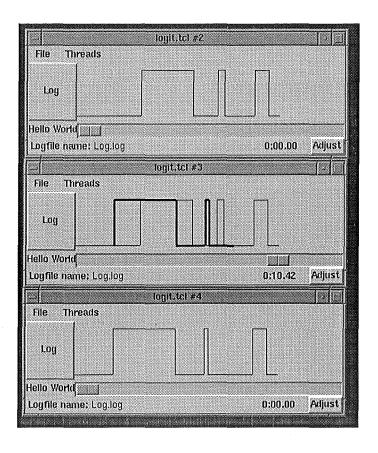


Figure 3: A thread being edited. top: Original thread, middle: After editing, bottom: After refreshing.

Pairs of threads could be analysed together, with ratios of events being taken, rather than the intersection. In a teacher–student interaction, you can measure the time spent by each talking, to gauge levels of understanding. It is probably more desirable for the student to have a low ratio of speaking compared to the teacher.

Analysis could be used to highlight unusual behaviour. If an unusual amount of hand-gesturing is observed then further analysis may be needed to determine why it is occurring.

One of the potentially most rewarding forms of analysis that could be carried out on threads is the search for patterns of behaviour. This takes advantage of their quantitative nature. It may even be possible to automate the search for such patterns, using a pattern matching algorithm. There are almost certainly patterns in a way an expert acts when speaking to another expert, when compared to how they talk to a learner. Searching for and

identifying such patterns could also come in useful for distinguishing unusual behaviour. These possible findings could be very useful.

5.3 Group awareness

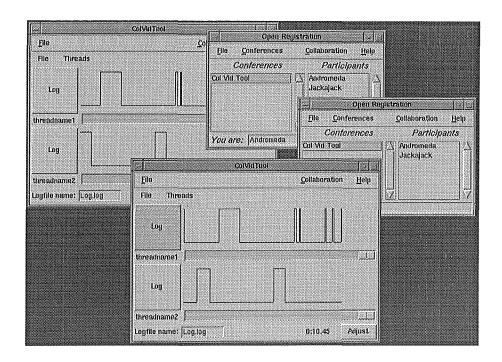


Figure 4: The full collaborative Mawd

Group awareness is essential for this tool, as this makes it unique. Evaluation of systems and interactions is done collaboratively, and so this tool must also support collaboration.

A significant advantage of a group logging threads together is that it cuts down the time needed in logging. A single user would have to watch the video segment as many times as there are actions to record, this will be time-consuming³. Multiple users need fewer sessions to log the same data.

Groups working together allow a greater flow of ideas. Deciding what actions to measure in threads needs much thought. There is less likely-hood that something is missed, and ideas can be shared, with more points of view.

The implementation of the groupware system is essentially the same as the multi-threaded, single user interface. Using GroupKit (Roseman & Green-

³Not to mention boring, with errors from loss of concentration!

berg, 1992) it was fairly trivial to allow collaborative use of threads. Initially the groupware design was to allow only one thread to be manipulated by each user at a time. This quickly proved to be a bad idea, as it made the use of the tool clumsy, and it forced roles on the users, when it is much easier for the user to assign roles, rather than having them forced upon them, especially since there roles are likely to be changing frequently.

6 Discussion and Future Work

Mawd is intended for the following styles of use:

- + It would be best used to analyse parts of an interaction that have been found to need a more thorough analysis. It can be used to gain a detailed description of what has occurred.
- + Mawd will make it more realistic to create longer threads than currently. The analysis done by Ishii et al (1992) of 140 seconds in duration could easily be extendible. This makes the results gathered are more suitable for statistical manipulation and analysis, rather than an ad-hoc indicator as used by Ishii et al (1992). It is unrealistic to document whole tapes with the tool, and this is not what was intended.
- + It should be used in conjunction with other tools. In particular it would be suitable for incorporation in a system like that outlined by Trigg (1989). In this paper, the system allows more than one form of annotating video. Threads could be added to augment or replace the current event-based representations that are used. Mawd should be used in conjunction with other forms of analysis.

The output is in a simple form, and the interface is also simple. This means that anyone that knows what they should be observing is able to use the tool. Other tools, such as EVA (Mackay, 1989) and that described by the Workplace project (Trigg, 1989) require more skill, and knowledge about what needs to be recorded, or annotated. Thus Mawd could be used by non-analysts to save time.

There are many additions and changes that could be made to Mawd.

- Editing. To edit a section of thread, it must be re-logged. Directly manipulating the thread would be a more useful option in many cases. Being able to move the thread directly would simplify many editing tasks. An undo function for editing would also be useful.
- Video synchronisation. The current methods of stopping and starting the logging bare no relation to that of the video. Something simple like a countdown before logging starts would make synchronising the two easier. The best option would be having the video-tool and video together on a window, and linked, so that starting the video starts logging. This would require a higher-powered multi-media machine.

- Representations. Currently there is just one representation of a thread, which is not immediately obvious as to its meaning. More semantically recognisable representations would be better. A hand to represent a hand-gesture, or eye for eye-movement. Even other forms of line, like a dashed one, instead of the current "snake", would allow threads to be overlayed.
- Threads. Currently threads have binary values. It could be possible to have analogue (continuous) values, or multiple-valued threads. Although this adds to the complexity, such logs would allow sophisticated analysis.
- Testing. The tool should be tested by conducting a basic collaborative logging task. This could be done recursively, videoing the testing, and analysing it with the tool.

7 Conclusion

Mawd is a collaborative video analysis tool that tackles three issues in logging video:

- 1. It is simple to use. The concentration is focused on observing interactions, not in using the interface.
- 2. It gives quantitative results. This is a unique feature in video analysis tools.
- 3. It is collaborative. Analysis is done collaboratively, and so the tool can be used collaboratively.

This report details the implementation of a collaborative video analysis tool. It summarises the role of evaluation in HCI and CSCW. It details video's attributes, that make it suitable for use in evaluation and it surveys some current video analysis tools.

8 Acknowledgements

Thanks to my Supervisor, Dr. Andy Cockburn, for his ruthless slashing of my text, endless supply of books and his constant enthusiasm. My collegues in the Honours' Room, for ensuring I didn't get a chance to enter the Realms of Sanity. Pepsico Inc., for in two words: "Mountain Dew".

References

- Allen, C. 1989. The use of video in organizational studies. SIGCHI bulletin, 21(2), 115–117.
- Allen, C., Ballman, D., Begg, V., Miller-Jacobs, H., Muller, M., Nielsen, J., & Spool, J. 1993. User Involvement In the Design Process: Why, Where and How? *INTERCHI Conferece Proceedings*, Apr., 251–254.
- Bentley, R., Hughes, J., Randall, D., Rodden, T., Sawyer, P., Shapiro, D., & Sommerville, I. 1992. Ethnographically-informed systems design for air traffic control. *Proceedings of the 1992 ACM Conference on Computer Supported Cooperative Work* October 31 to November 4 1992. Toronto, Canada., Nov., 123–129.
- Brun-Cottan, F., & Wall, P. 1995. Using Video to Re-Present the User. Communications of the ACM, 38(5), 61-70.
- Carter, K., & Anderson, B. 1989. Can Video Research Escape the Technology? Some reflections on the problems and possibilities of Audio Visual research. SIGCHI bulletin, 21(2), 112–114.
- Dix, A., Finlay, J., Abowd, G., & Beale, R. 1993. Human-Computer Interaction. Prentice Hall.
- Fetterman, D. 1989. Ethnography: Step by Step. Applied Social research Methods, vol. 17. Sage Publications.
- Hammersley, M, & Atkinson, P. 1983. Ethnography: Principles in Practice. Tavistock Publications.
- Harrison, B., Mantei, M., Beirne, G., & Narine, T. 1994. Communicating About Communicating: Cross-Disciplinary Design of a Media Space Interface. *Proceedings of CHI'94 Conference on Human Factors in Computing Systems* Boston, April 24–28 1994, Apr., 131–137.
- Heath, C., & Luff, P. 1982. Collaboration and Control: Crisis management and Multimedia Technology in London Underground Line Control Rooms. 1(1-2), 69-94.
- Horton, M., Elwart-Keys, M., & Kass, R. 1989. Video as an enabling technology for computer supported cooperative work. SIGCHI, 21(2), 96–99.

- Hughes, J., Randall, D., & Shapiro, D. 1992. Faltering from Ethnography to design. *Proceedings of the 1992 ACM Conference on Computer Supported Cooperative Work* October 31 to November 4 1992. Toronto, Canada., Nov., 115–122.
- Hughes, J., King, V., Rodden, T., & Andersen, H. 1994. Moving Out from the Control Room: Ethnography in System Design. Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work October 22 to 26. Chapel Hill, North Carolina., Oct., 429-439.
- Ishii, H., & Kobayashi, M. 1992. ClearBoard: A Seamless Medium for Shared Drawing and Conversation with Eye Contact. *Proceedings of CHI'92 Conference on Human Factors in Computing Systems* Monterey, May 3–7 1992, May, 525–532.
- Ishii, H., Kobayashi, M., & Grudin, J. 1992. Integration of Inter-Personal Space and Shared Workspace: ClearBoard Design and Experiments. Proceedings of the 1992 ACM Conference on Computer Supported Cooperative Work October 31 to November 4 1992. Toronto, Canada., Nov., 33–42.
- Mackay, W. 1989. EVA: An experimental video annotator for symbolic analysis of video data. SIGCHI bulliten, 21(2), 68-71.
- Monk, A., Nardi, B., Gilbert, N., Mantei, M., & McCarthy, J. 1993. Mixing oil and water? Ethnography versus Experimental Psychology in the study of computer-mediated communication. *Interchi Conference Proceedings*, Apr., 3–6.
- Neal, L. 1989. The use of video in Empirical research. SIGCHI bulletin, 21(2), 100–102.
- Nielsen, J. 1993. Usability Engineering. Academic Press.
- Rieman, J., Lewis, C., Young, R., & Polson, P. 1994. "Why Is a Raven Like a Writing Desk?" Lessons in Interface Consistency and Analogical Reasoning from Two Cognitive Architectures. *Proceedings of CHI'94 Conference on Human Factors in Computing Systems* Boston, April 24–28 1994, Apr., 438–444.
- Rose, D. 1990. Living the Ethnographic Life. Qualitative Research Methods, vol. 23. Sage publications.

- Roseman, M., & Greenberg, S. 1992. GroupKit: A Groupware Toolkit for Building Real-Time Conferencing Applications. *Proceedings of the 1992 ACM Conference on Computer Supported Cooperative Work* October 31 to November 4 1992. Toronto, Canada., Nov., 43–50.
- Roske-Hofstrand, R. J. 1989. Video in applied cognitive research for human-centred design. SIGCHI bulletin, 21(2), 75–77.
- Shapiro, D. 1994. The limits of Ethnography: Combining Social Sciences for CSCW. CSCW94, Oct., 417–428.
- Suchman, L., & Wynn, E. 1984. Procedures and Problems in the Office. Office Technology and People, 2, 133–154.
- Tang, J. 1991. Findings from observational studies of collaborative work. *International Journal of Man-Machine Studies*, **34**, 251–259.
- Trigg, R. H. 1989. Computer Support for transcribing recorded activity. SIGCHI bulletin, 21(2), 72–74.
- Twidale, M., Randall, D., & Bentley, R. 1994. Situated Evaluation for Cooperative Systems. *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work* October 22 to 26. Chapel Hill, North Carolina., Oct., 441–452.
- Weber, K., & Poon, A. 1994. Marquee: A Tool for Real-Time Vidwo Logging. Proceedings of CHI'94 Conference on Human Factors in Computing Systems Boston, April 24–28 1994, Apr., 58–64.
- Weir, L. 1994. Digital Video For Time Based Analysis Systems. M.Phil. thesis, University of Canterbury.