DTU Library

Multi-modal recording and analysis of interaction among operators and work systems

Andersen, H.B.; Hansen, John Paulin

Publication date:
1996

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Andersen, H. B., & Hansen, J. P. (1996). *Multi-modal recording and analysis of interaction among operators and work systems*. Denmark. Forskningscenter Risoe. Risoe-R No. 939(EN)

Risø-R-939(EN)

RISØ

Multi-Modal Recording and Analysis of Interaction among Operators and Work Systems

Henning Boje Andersen & John Paulin Hansen

Risø National Laboratory, Roskilde, Denmark
December 1996

Abstract This paper describes a new PC- based system, MULTIMO, which enables human factors specialists (a) to analyze and identify correlations among operators' behaviours and work environment characteristics - e.g., correlations between alarms, detection times and subjects' activation of controls - and (b) to compare such correlations across different individuals, tasks, scenarios or training backgrounds as well as different types of control interfaces and task procedures. The system facilitates the scoring of the multiple, synchronous a/v recordings and analysis and multimedia visualization of results.

ISBN 87-550-2241-3
ISSN 0106-2840

Grafisk Service, Risø, 1996

Contents

1. Introduction[1]

Modern technical, complex work environments typically involve one or several subjects (ie, pilots, process operators, anaesthetists, etc.) interacting with one or several control panels and other parts of their work situation. There are several problems in collecting and analysing data in order to arrive at informed judgements about operator performance and various work environment characteristics, e.g., alarm configurations.

Synchronous audio and visual recordings of operators coping with various tasks have been made for years, of course. Yet, only fairly recently has it become practically feasible for a human factors specialist to compare larger bodies of transcriptions of verbal output with synchronous video sequences and in fact, operate with ease a VRT via a PC. However, in order to assess operator behaviour in, say, an aircraft simulator or a NPP control room, it is often desirable to have a record of several of the behavioural *modalities* of the subjects. For instance, it may not be sufficient to have available audio recording of verbal output in addition to the log of selected work environment (simulator) events and log of operators' activation of controls. That is, it may be desirable to have available as well recordings of eye movements and fixations - and perhaps body movements and posture may be essential in order to conduct a cognitive analysis of operator behaviour. Moreover, it is not enough that such logs and recordings be available separately; what is needed is a *synchronous* replay facility that can access display any combination of logs and recordings and supporting "dynamic links" between items or scenes of different logs/recordings. As will be explained below, the design of MULTIMO goes some way towards satisfying this requirement.

Second, human factors analysts are well aware that it has been inherently time-consuming and error prone, using traditional observational research methods, to track the fast flow of information in safety critical real time environments (cf. Sanderson and Fisher, 1994). This has put strong limitations on the number of cases analysed and on the contextual depth explored. It may take several hours to analyse only the attention allocation of an operator during ten minutes of process control. In addition, the subject's verbal communication and activation of controls have to be analysed as well, and this will inevitably add several hours to the task. Often the ratio of analysis time to observational session time ends up being 50 to 1 or even more, and for all practical research purposes this is unacceptable. One of the main goals of developing MULTIMO has been to reduce this ratio substantially while supporting the capability of displaying and scoring several tracks simultaneously.[2]

2. MULTIMO functionalities

MULTIMO supports the research process at four stages: (1) recording and logging, (2) categorisation and scoring, (3) statistical analysis and (4) visualisation and reporting.

---

[1] This report is a modified and somewhat expanded version of Andersen and Hansen 1995. Reprint of the paper with the same title appearing in: *Proceedings of Fifth International Conference on Human-Machine Interaction and Artificial Intelligence in Aerospace* - HMI-AI-AS'95, Toulouse, France, September 27-29 1995

[2] We refer to the superb and wide-ranging overview paper by Sanderson et al. 1994 and to the special journal issue to which Sanderson & Fisher 1994a forms the introduction.

2.1 Recording and logging

The types of behaviours that are candidates for recording include

- eye movements of one or two subjects by means of an eye tracking system (ASL type);
- hand movements of one or two subjects, by means of two separate video cameras and simple, coloured gloves or wristbands;
- separate views of the total work scene including the control panels by means of typically one or two - but in certain cases several - video cameras;
- stereo recordings of speech and sounds of the work environment;
- the activation of controls (when and where switches and dials are turned, their position, etc.) as well as selected work environment parameters (e.g. alarms and process state variables). Since the selection of such parameters is highly dependent on the technical interfaces of the equipment of the specific work environment, the generic requirement to a log is that logged events be time stamped in synchrony with the a/v recordings (see below about time tags).

The different types of recordings are strictly synchronised with a precision of half a frame-to-frame correspondence (1/50 second) when two or several VCRs are used. The log of activation of controls and the log of environmental parameter values receive the same time tags as that which is inserted into the audio/video track(s), but the precision of the log is obviously no greater than the timing precision and resolution of the simulator and equipment that produce the log. Furthermore, tokens of 12 individual time tags can be added to the recordings - either while the observation session is running or afterwards. This feature allows a human factors analyst to insert tags indicating *tokens of* 12 different *types* of events or intervals of interest - single point events need a single tag whereas intervals are indicated by toggle button, the first tag marking the start and the second the end of a given interval. Tags are inserted by the use of a standard remote control device. The 12 different tags can be predefined as determining a point event or the start/end of an interval - i.e., they are assigned as simple pulse indicators or as on/off toggles; moreover, they can be predefined in groups as being mutually exclusive within that group (e.g., types of speech acts by a single subject). In well-structured and controlled experiments, this enables the human factor specialist (or several specialists) to do most of the scoring "on the fly". Alternatively, the tags may be used to collect several independent, possibly simultaneous, expert assessment of the performance observed.

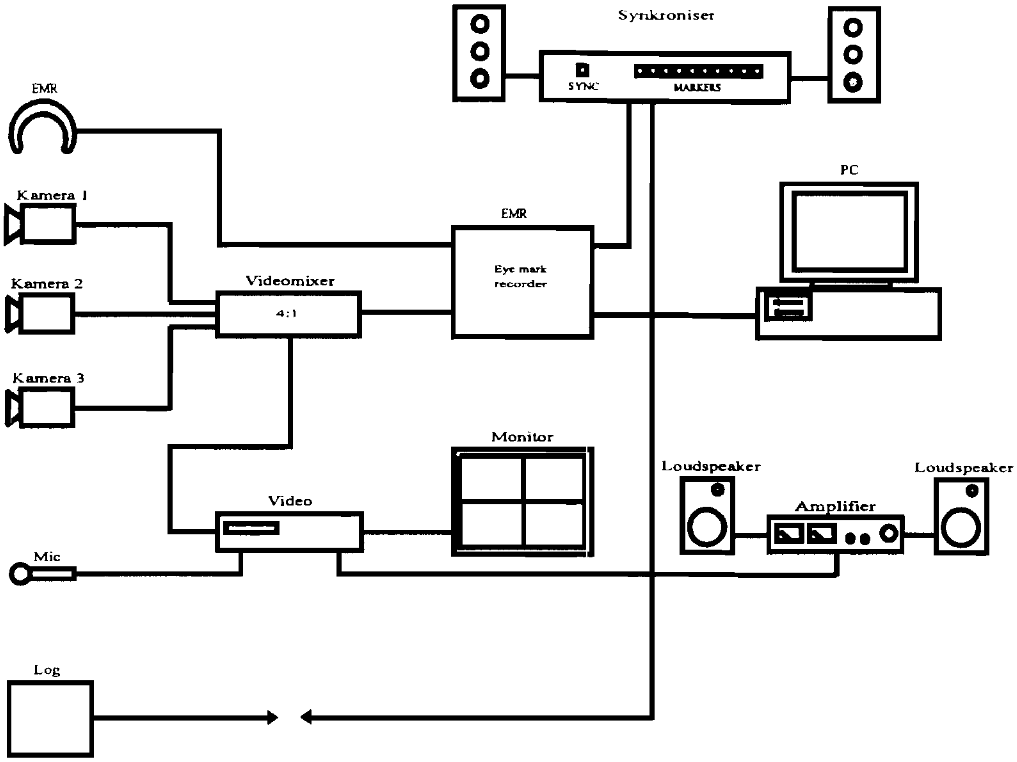

Figure 1: Block diagram of the basic MULTIMO configuration with one eye mark recorder and a single VCR

Figure 1 gives an overview of the basic recording configuration. It consists of 3 cameras plus one eye mark recorder. The four video signals are combined in a quad-unit, mixing them into a single video signal to be recorded on one VCR together with the audio signals. The synchronisation is done with an ASL eye tracker as the central unit, using its digital data collection channels. When needed, it is possible to add one or several additional VCRs.

When the ASL eye tracker is not included in a concrete configuration, its recording mode is emulated by a simple software program in order to get the signals from the synchroniser recorded. The synchroniser sends a signal to the database of the eye tracker, to the log of the simulator, and to two flashing lights located in view of the 3 video cameras, when the human factor specialist presses a fire button on its front panel. Alternatively, it may fire automatically on receiving a specified analog or digital signal emitted from the simulator. If no direct communication with the simulator is possible, the synchronisation may be done post hoc on basis of an event, e.g. an alarm, with can be identified in the simulator and can be seen or heard in the video recordings.
The system concept is modular allowing for another eye mark recorder to be added and, as mentioned above, several VCRs that will run in parallel, each recording the video signals from one or four cameras.

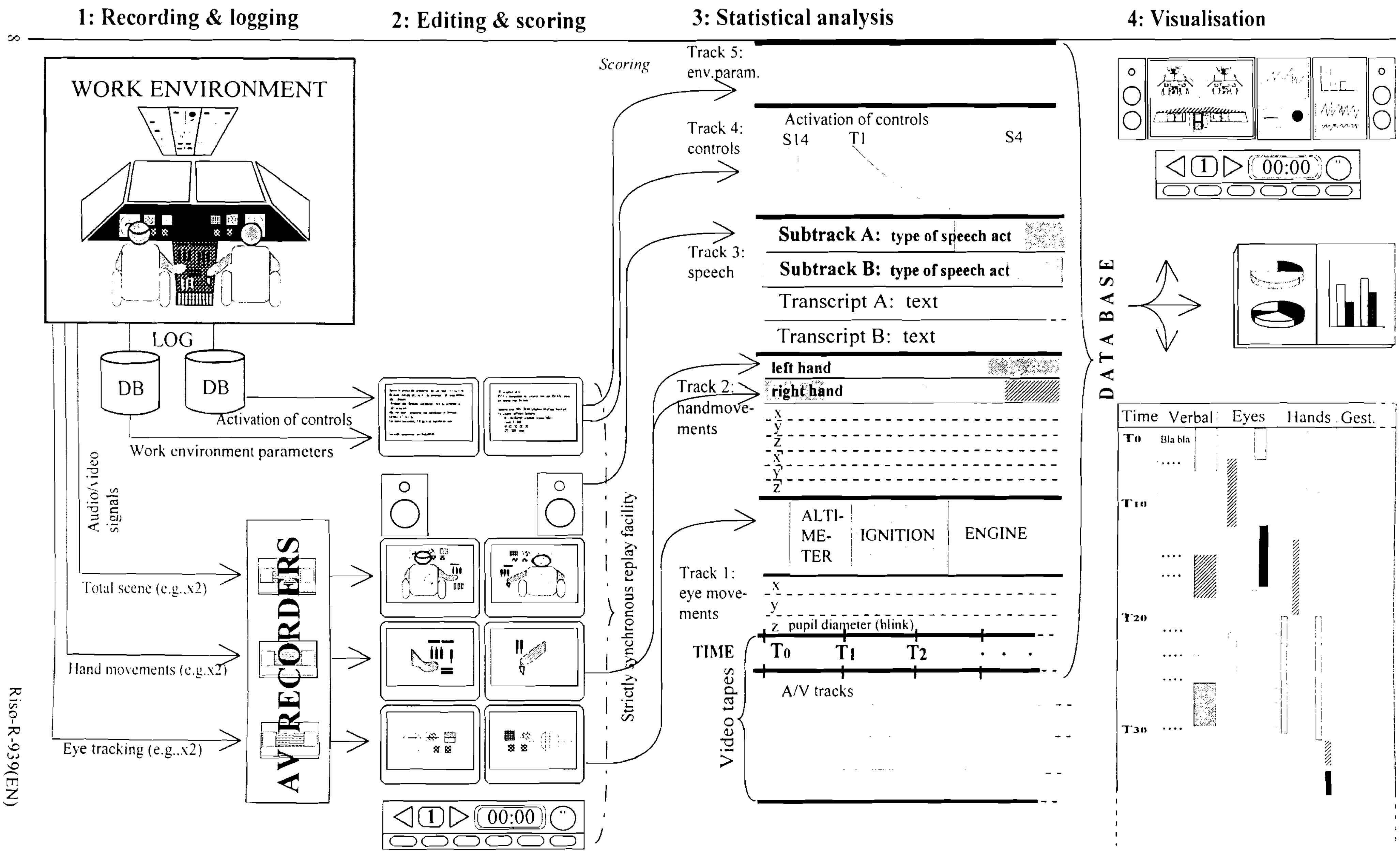

Figure 2: An overview of how the four phases during recording and analysis of operator behaviour are supported by the MULTIMO System

2.2 Categorization and scoring

In the second phase, recordings and logs will be analyzed and reformatted yielding a temporal database file (i.e., a set of records in a temporal database). The purpose of creating the temporal datafile is to allow human factors analysts (a) to address queries concerning the type of correlations mentioned above, e.g., relations between alarms detection times compared across different operators, tasks or work environment configurations; (b) to combine replay of the a/v tracks and logs with synchronous presentation (graphical, character-based or charts) of data.

In order to support the queries the temporal datafile is created containing selected information from the a/v tracks as well as the logs of control activation and environment parameters. The file contains, for each of the "modalities" that are recorded, a "temporally ordered track" consisting of selected information from the a/v recordings as well as the logs (confer Figure 2).

There are two kinds of information derived from the a/v recordings that go into the temporally ordered tracks of the temporal datafile. For a given datafile track, there will be, first, parameter values that are to be taken automatically from the corresponding data logs (i.e. from the simulator and/or from the eye movement data file) and second, category scores done manually by an analyst. For instance, from the eye tracker, for each position (x & y) and pupil diameter (mm), there is a automatic transfer of values, for every 50 milliseconds, to three subtracks of the"eye movement track" (confer Figure 2, column 3).

The manually scored categories are typically a most important part of the temporal database. In order to create these categories it is necessary to have available the synchronous replay facility described in the following paragraphs. The categories are meant to reflect"areas of interest" or focal points for each of the modalities concerned. Formally, they are non-overlapping intervals that, for a given modality, instantiate any one among a limited number of categories. For instance, for the speech track the categories will be types of speech acts or conversational topics; for the eye movement track the categories will be items in the work environment fixated by the subject; and for the hand movement track the categories will (typically) be the controls touched by the subject's hands. Accordingly, the specific subtrack devoted to category scoring within a given track we call "areas of interest band".

While it will no doubt be possible to go some way towards automating the category scoring, we do not aim at including this for the first version of the equipment. Instead, we intend to provide support for a convenient manual scoring. Thus, the replay facility is able to provide strictly synchronous replay of all the a/v recordings (and any subset of these) as well as the annotated log files (activation of controls; work environment parameters). The output of the a/v recordings and the log files can be shown on separate monitors, on separate windows of a PC screen (if a high performance video board is included or on a single TV monitor with 4 split screens. The audio output is going to an amplifier and to stereo speakers.

The replay facility includes all the functions of an advanced VCR control. The control can be operated by icon clicks on a graphic module included in the database display or by hot keys. It supports single, simultaneous frame steps and it allows access to any time tagged sequence or frame. In a sense, the database itself is a VCR control: Selecting any single (or excerpts of) data in the log or in the manual scorings from the database automatically brings the VCR to the position and ready replay.

In a so called "User column" of the database it is possible to enter free text supported by simple editing functions. This column may be used to transcribe verbal protocols or attach more comprehensive comments to the recordings reviewed. However, transcription of verbal protocols is indeed one of the most trivial and time consuming parts of a task analysis. Further more, lots of important information in, e.g., tone of voice and pausing, get lost during the transcriptions. Therefore, by the use of a sound card, it is possible to sample relevant parts of the audio protocol and store it as digitized audio notes in the database. In principle, only the capacity of the systems hard disk puts a limit to the amount of communication included.

Finally, a most important scoring support is provided by a module in which up to 25 categories can be labeled with a 20 character (maximum) text string. Each of the categories is assigned a button. When a particular button is pressed during the review, its category label is written into the database and is "time stamped" accordingly. For analysis with high demands on time accuracy the replay may be done at very slow speed and for quick and approximate scorings it may run at fast forward.

The 25 buttons can be located anywhere at the screen and on top of a bit-map picture. For instance, the picture can be a grabbed video image of the work scene or it can be a screen dump of a particular display often attended by the operator. To score, say, the hand movements on system controls or the eye fixations on panels, the human factor specialist locates his buttons on top of the areas of interest in this scene. Then he or she goes through the recording and simply moves the courser onto the various areas touched- or fixated - by the operator. This is to insure a direct mapping from the recording medium to the scoring medium, which - in principle - allows people (e.g. research assistants) without knowledge of the domain categories behind the area of interest to do the scoring. And, we may add, to do the scoring in a smooth"direct manipulation" fashion.

Four columns of the database have been reserved for this type of scoring, allowing categories at four dimensions to be added successively. Alternatively, two human factor specialists may conduct scoring of the same episode independently with identical sets of categories to control for inter-subjective differences in the scoring. In this case, it is possible to hide the column used by a previous analyst.

2.3 Statistical analysis

The temporal database has a format that is compatible with standard statistical software programs running on personal computers (PCs or Macintoshes). The statistical software to be used is external to the Multimo technology (though of course essential).

The analysis phase can take one of two forms: first, there will be the comparative study that seeks to assess differential behavioral outcomes across two or more distinct sets of conditions. (For instance, studies aiming at identifying differences in subjects' behavior and performance as related to two (or several) kinds of procedures, training, control interfaces, tasks, subject profiles). Second, there will be specific studies aiming at uncovering the details of why, say, a given control interface leads to diminished performance; or it may seek to ascertain the extent to which, say, unusual symptoms are being noticed by operators working in naturalistic and normal circumstances.[3]

[3] The current applications of Multimo include investigations of alarm management in anaesthesia, transfer of learning with respect to touchscreen-based part task trainers, comparison of interfaces to ship pilot computer support facilities.

Corresponding to this rough division into two types of studies, the types of questions that will be investigated will concern, on the one hand, comparisons and correlations/differences across different experimental settings and, on the other, concrete relations within a single type of setting. The requirements to databases used for further analysis are therefore (1) that it must support queries which are "multi-dimensional" - i.e. queries which refer to waiting times between and among *several* subtracks (categories within the "areas of interest bands"); and (2) that it must be compatible with standard statistics software.

Statistical analysis of the eye tracking data has certain demands not directly supported by standard software. Therefore we have developed a stand-alone-tool for eye data analysis, called "Eye Track 2", which runs under the same Windows platform as MULTIMO and sharing important functionalities with it. Eye movement patterns generated from position data are highly dependent on the defined criteria for a fixation in terms of time (e.g. from 100 to 250 milliseconds) and in terms of maximum spatial deviation within a single fixation. There is no general agreement on these criteria in the literature, as they vary with the type of individual and type of task analyzed. As a consequence, the criteria must be defined to match the pattern observed in the video recordings of the eye movement. In order to do so, the analyst needs to adjust them on basis on direct comparisons with the video. This is where the VCR control functions of MULTIMO comes in handy: By clicking on a sequence of fixations being displayed in Eye Track 2, the analyst can see the corresponding sequence of the video recording, adjust his criteria and have a new fixation pattern generated immediately.

The areas of interest defined as buttons in MULTIMO may be imported directly into Eye Track 2. This allows the analyst to generate summary statistics of, for example, the amount of attention being paid to certain areas and to compare this with the outcome of the manual scoring. It may be necessary to adjust the eye movement pattern manually to compensate for head movements before the comparison can be done. Eye Track 2 supports these adjustments by means of rotation, dragging, stretching, etc. and with the possibility of having a grabbed video scene image from MULTIMO as an underlying frame of reference. In addition, Eye Track 2 provides several options for analysis and display of eye data, but it is beyond the scope of this paper to describe them in detail.

2.4 Visualization

When the scoring and statistical analysis are finished they can all be integrated into one multimedia document. The user can e.g. see all instances of a "human error" scoring on video, listen to the audio clips selected, see snapshots of gabbed video pictures from particular situations, see the statistical and graphical documents created by Eye Track 2 or by other modern spreadsheets like Excel, Quattro, charting tools contained in statistics software, etc.

The visualization part of MULTIMO, containing the excerpts and annotations made by the analyst, is also temporally structured. This allows the analyst to make precise reference to the visualizations available in an accompanying text report.

References

Andersen, A.W., (1995): Multimo Brugermanual [Multimo User Manual]. Riso National Laboratory, Electronics Department

Andersen, H.B. & Hansen, JP (1995): Multi-modal recording and analysis of operator data. In *Proceedings of Fifth International Conference on Human-Machine Interaction and Artificial Intelligence in Aerospace - HMI-AI-AS'95*, Toulouse, France, September 27-29 1995

Sanderson, P.M. & Fisher, C. (1994a): Introduction to special issue on exploratory sequential data analysis. *Human -Computer Interaction*, 9, 247-50

Sanderson, P.M. & Fisher, C. (1994b): Exploratory sequential data analysis: Foundations. *Human -Computer Interaction*, 9, 251-318

Sanderson, P.M., Scott, J., Johnston, T., Mainzer, J., Watanabe, L. & James, J. (1994): MacSHAPA and the enterprise of exploratory sequential data analysis (ESDA). *International Journal of Human-Computer Studies*, 41, 633-681.

Bibliographic Data Sheet **Risø-R-939(EN)**

Title and author(s)

Multi-Modal Recording And Analysis Of Interaction Among Operators And Work Systems

Henning Boje Andersen and John Paulin Hansen

ISBN	ISSN
87-550-2241-3	0106-2840

Dept. or group	Date
Systems Analysis Department	December 1996

Group's own reg. number(s)	Project/contract no.

Pages	Tables	Illustrations	References
12			

Abstract

This paper describes a new PC- based system, MULTIMO, which enables human factors specialists (a) to analyze and identify correlations among operators' behaviours and work environment characteristics - e.g., correlations between alarms, detection times and subjects' activation of controls - and (b) to compare such correlations across different individuals, tasks, scenarios or training backgrounds as well as different types of control interfaces and task procedures. The system facilitates the scoring of the multiple, synchronous a/v recordings and analysis and multimedia visualization of results.

Descriptors INIS/EDB

Available on request from

Information Service Department, Risø National Laboratory
(Afdelingen for Informationsservice, Forskningscenter Risø)
P.O. Box 49
DK-4000 Roskilde, Denmark
Telephone (+45) 4677 4677, ext. 4004/4005
Telex 43 116, Telefax (+45) 4675 5627

RISØ

Objective

Risø's objective is to provide society and industry with new opportunities for development in three main areas:

- *Energy technology and energy planning*
- *Environmental aspects of energy, industrial and agricultural production*
- *Materials and measuring techniques for industry*

In addition, Risø advises the authorities on nuclear issues.

Research profile

Risø's research is strategic, which means that it is long-term and directed toward areas which technological solutions are called for, whether in Denmark or globally. The research takes place within 11 programme areas:

- *Wind energy*
- *Energy materials and energy technologies for the future*
- *Energy planning*
- *Environmental impact of atmospheric processes*
- *Processes and cycling of matter in ecosystems*
- *Industrial safety*
- *Environmental aspects of agricultural production*
- *Nuclear safety and radiation protection*
- *Structural materials*
- *Materials with special physical and chemical properties*
- *Optical measurement techniques and information processing*

Transfer of Knowledge

Risø's research results are transferred to industry and authorities through:

- *Co-operation on research*
- *Co-operation in R&D consortia*
- *R&D clubs and exchange of researchers*
- *Centre for Advanced Technology*
- *Patenting and licencing activities*

And to the world of science through:

- *Publication activities*
- *Network co-operation*
- *PhD education and post docs*

Key Figures

Risø has a staff of more than 900, including more than 300 researchers and 100 PhD students and post docs. Risø's 1996 budget totals DKK 471 m, of which 45 % come from research programmes and commercial contracts, while the remainder is covered by government appropriations.

Risø-R-939(EN)
ISBN 87-550-2241-3
ISSN 0106-2840

Available on request from:
Information Service Department
Risø National Laboratory
P.O. Box 49, DK-4000 Roskilde, Denmark
Phone +45 46 77 46 77, ext. 4004/4005
Telex 43116, Fax +45 46 75 56 27
http://www.risoe.dk
e-mail: risoe@risoe.dk