in: T. Grechenig & M. Tscheligi (eds.) Human Computer Interaction. (Lecture Notes in Computer Science 733, pp. 129-140), Berlin: Springer.

Locating the primary attention focus of the user

Matthias Rauterberg & Christian Cachin

Work and Organisational Psychology Unit, Swiss Federal Institute of Technology (ETH)
Nelkenstrasse 11, CH-8092 ZURICH
rauterberg@ezrz1.vmsmail.ethz.ch

Abstract. First, a signal detection experiment was carried out to estimate the maximal distance between the primary attention focus of users and the screen position of visual feedback (e.g. messages). The results indicate that the maximal distance between the primary attention focus and the position of visual feedback should not exceed 3". Second, to pinpoint the location of the primary attention focus we carried out an eye movement recording experiment. The results indicate that if the task solving process requires mouse operations and the visual feedback of the results of these mouse operations appears close to the mouse cursor, then the visual focus and the mouse cursor position on the screen are highly correlated: between 76% and 95% correspondence.

Keywords: primary attention focus, visual focus, mouse control, visual feedback, user interface, design.

1 Introduction

One important problem in interface design is making appropriate design decisions regarding the positioning of visual feedback on the screen (e.g. messages, alert boxes, hints, pop up menus, icons, etc.). While highlighting techniques can aid the user in locating important messages, it is not always possible to predict what may be important to the user at a given time.

The traditional solution is a mask layout that allows the user to" easily" find any of the information on it by adopting a consistent format for all masks of a character user interface (CUI). Following the guidelines of [1] and [2] the screen layout looks as follows:

(1) the top line contains the screen title, time of day, and product name;
(2) the second line contains a command input field;
(3) the third line contains a list of the commands that are currently valid;
(4) the middle part of the screen contains the work area (e.g. output area for menus, data lists, data entry forms, etc.);
(5) messages appears in the bottom two lines.

This mask layout has been empirically "verified" (e.g. [10], [11]). One problem with these kinds of empirical studies is the experimental control group condition: both studies compared structured and consistent mask layouts on the one hand with accidentally arranged mask layouts on the other hand. Under these specific circumstances it is very probable, that each kind of 'structured' conditions is better than a total 'unstructured' condition.

The following design decision tries to overcome this problem in the context of the design of graphic user interfaces (GUIs): important messages appears in the centre of the screen (see [8]). This solution minimises the distance between the unknown locus of the primary attention focus of the user and the locus of the message on the screen.

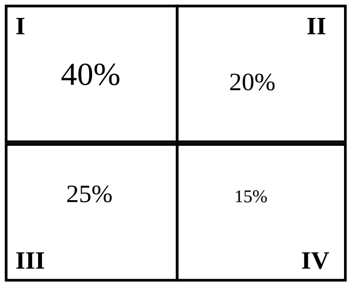

Figure 1. The relative ratios of the user's visual focus looking expectantly on one of the four quadrants of a dark and unstructured screen [9].

What is an *optimal* screen layout is till now the open and unanswered question. Where is the *best* place to put messages on the screen? How far away from the primary attention focus should the message be? To answer these questions, first we carried out a 'signal detection experiment'. From the literature [9] we know that the main area of expectation (40%) is in the left upper quadrant I of a dark and unstructured screen. The visual focus is further on shared by quadrant III (25%) and quadrant II (20%). In 15% of observation time the attention focus is in quadrant IV (see Figure 1). First, we carried out a signal detection experiment to estimate the maximal distance between the primary attention focus of users and the screen position of visual feedback (e.g. messages) regarding to the four screen quadrants. Second, we investigated the eye movements to pinpoint the actual location of the primary attention focus of the user on the screen regarding to different task's types.

2 Signal Detection Experiment

One major determinant of both the context and the quality of perception is *attention*. Foley and Moray [4] present a list that at least the following variables play a role in controlling dynamic visual attention:

(1) the rate at which the display varies: the greater the bandwidth, the more frequently is the display sampled;

(2) the value of the information: the more the information is worth, the more frequently the display is sampled;

(3) the cost of observation: the more costly an observation, the less frequently is the display sampled;

(4) forgetting: as time elapses since the last observation, the user becomes less certain of the value of the observed information even if it varies only slightly or not at all;

(5) the coupling between displays.

We proved aspect (2) and (3) with the dimension 'signal type' as a value indicator and with the dimension 'distance' as a cost indicator in this signal detection experiment.

2.1 Subjects

Eleven women and 8 men took part in the experiment (mean age: 33 ± 14 years). 12 subjects were students of computer science at the ETH.

2.2 Tasks

Subjects were introduced to solve two tasks: the primary task was 'counting a small set of circles' (6 – 10 elements) presented in one of the four screen corners (visible: 1000 ms; area: 2.5" x 2.0"), and the secondary task was 'detecting an X' (see Figure 2).

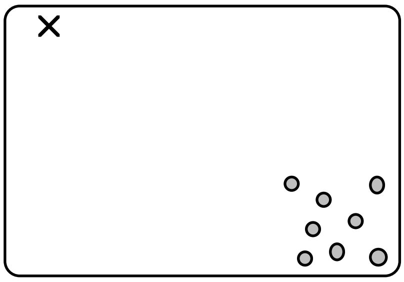

Figure 2. An example screen with 8 circles in quadrant IV and an X signal in quadrant I in the 'large distance' condition.

To control the perception of the semantic of the signal we presented two different signs: an X and a square of the same size. One of the two signals appeared after 500 ms of circle presentation time and was visible for the duration of 500 ms along with the circles (see Figure 3). The subjects were requested to respond only to the X signal. The attention focus of the users was controlled by the primary task.

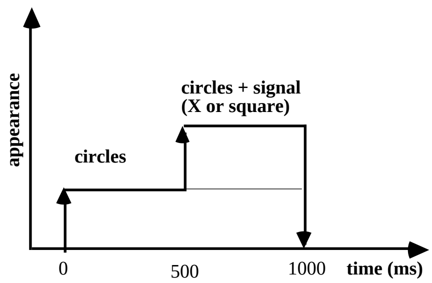

Figure 3. The time structure of the signal presentation.

2.3 Procedure

We have a nested test design, because the factor A and C are only possible for the signal type 'X' and 'square':

Factor A 'distance' (D) between primary attention focus and signal (inside 0", small 0"-3", medium 3"-6", large 6"-9") – the factor A was only varied for the signal type 'square sign' and 'X sign';
Factor B 'signal type' (4 masks with 'no signal'; 16 masks with a 'square' sign [8 mm x 8 mm]; 16 masks with an 'X' sign [8 mm x 8 mm]);
Factor C 'screen quadrant' of circle output (top left [tl], top right [tr], bottom left [bl], bottom right [br]).
Each user saw 36 different masks (individual session). Overall we got 684 different data points (19*4 = 76; 19*(16+16) = 608; 76 + 608 = 684).

2.4 Material

We ran the experiment on an IBM compatible PC (Olivetti M386) with a colour screen (VGA, 14"). A special program was developed in MacMETHModula to present the signals on the screen. Each subject saw 16 masks with circles and an X sign, 16 masks with circles and a square sign, and 4 masks with circles only. We randomised the order of all masks and number of presented circles.

2.5 Measures

Three dependent measures are used in this study: (1) 'signal deviation', (2) 'error ratio', (3) 'circle deviation'. The signal detection table (see Table 1) is the basis of measure (1) and (2).

Table 1. The 'signal detection table' with the 'a', 'b', 'c' and 'd' as names of the four different cells.

		NO X SIGN (nothing or square)	X SIGN PRESENTED
answer of the subject	NO	a	b
	YES	c	d

The investigator protocolled the number of counted circles and the subjects' answer whether he or she had seen an X. The 'signal deviation' (XD) estimates the quality of the secondary task ('detection of an X'):

$$XD = |\#X_{detected} - \#X_{presented}| \quad (1)$$

XD is zero for all cases in cell 'a' and 'd'. The 'error ratio' (ER) is based on the 'signal detection table' (see Table 1).

$$ER = (b + c) / (a + d) * 100\% \quad (2)$$

Cell 'b' contains the number of 'overlooked signals'; cell 'c' contains the 'false alarms' and/or 'misinterpretations' of the 'square sign'.
The 'circle deviation' (CD) rate is a measure of the accuracy of the primary task ('counting of circles'):

$$CD = |\#CIRCLES_{counted} - \#CIRCLES_{presented}| * 100\% / \#CIRCLES_{presented} \quad (3)$$

CD is an independent measure of the absolute number of presented circles. With CD we can control the extent to which the users concentrate on the primary task solving process.

2.6 Results and Discussion

To analyse the data with a full factorial analysis of variance we computed the analysis only for masks with an X or a square sign (N=608). No detection errors ('false alarms') occurred at all in the signal type condition 'no signal' (N=76). We decided to exclude these data from the analysis of variance of XD. So, the results of XD are based on a full factorial analysis of variances.

The factor A 'distance' of the dependent variable XD is significant (N=152, $XD_{0''}$= 6%; N=152, $XD_{3''}$=11%; N=152, $XD_{6''}$= 27%; N=152, $XD_{9''}$= 43%; p≤.001, see Table 2). It is interesting to note that the factor B 'signal type' is not significant (N=304, XD_{x}= 23% 'overlooked'; and N=304, XD_{square}=20% 'misinterpreted'; p≤ .343, see Table 2). This result indicates that users have on average the same problems to detect an X independent of the semantic of the signal type.

Table 2. Results of the full factorial analysis of variance of the dependent variable XD (all masks of signal type 'no signal' are excluded).

Source	df	∑ of 2	means2	F test	p
A distance	3	12.847	4.282	29.001	.001
B signal type	1	.133	.133	.902	.343
C quadrant	3	.176	.059	.397	.755
A x B	3	1.308	.436	2.952	.032
A x C	9	1.410	.157	1.061	.390
B x C	3	.926	.309	2.090	.100
A x B x C	9	.923	.103	.694	.714
error	576	85.053	.148		

Table 3. Results of the full factorial analysis of variance of the dependent variable CD (all masks of signal type 'no signal' are excluded).

Source	df	∑ of 2	means2	F test	p
A distance	3	251.12	83.73	1.063	.364
B signal type	1	.75	.75	.010	.922
C quadrant	3	611.02	203.67	2.585	.052
A x B	3	74.46	24.82	.315	.815
A x C	9	898.18	99.80	1.267	.252
B x C	3	119.16	39.72	.504	.680
A x B x C	9	267.31	29.70	.377	.946
error	576	45387.	78.80		

The factor C 'screen quadrant' of the dependent variable XD is not significant, too (N=152, XD_{tl}=22%; N=152, XD_{tr}= 19%; N=152, XD_{br}= 24%; N=152, XD_{bl}=21%; p≤.755, see Table 2). The interaction term (A x B, see Table 2) is sig nificant (N=76, $XD_{0'',X}$= 4%; N=76, $XD_{0'',square}$= 8%; N=76, $XD_{3'',X}$=13%; N=76, $XD_{3'',square}$= 8%; N=76, $XD_{6'',X}$=24%; N=76, $XD_{6'',square}$=30%; N=76, $XD_{9'',X}$= 51%; N=76, $XD_{9'',square}$=34%; $p \leq .032$, see Table 2). This result is primarily caused by the difference between the average $XD_{9'',X}$=51% and $XD_{9'',square}$=34% in the large distance (9") condition. The user's have significantly more problems to detect an X sign in the large distance condition than to misinterpret a square sign as an X.

The significant effect 'quadrant' of the variable CD ($p \leq .052$, see Table 3) means that the primary task 'counting circles' is differently influenced by the secondary task 'detecting an X'. One can observe the greatest relative deviation between presented and counted circles (CD) in quadrant II and III (see Figure 4). In quadrant II the user must look 'left-down', in quadrant III he has to look 'right-up' to detect an X. Quadrant IV contains the lowest deviation of CD. It seems to be the easiest way for the user to look 'left-up' to detect a small signal (e.g. see Figure 2).

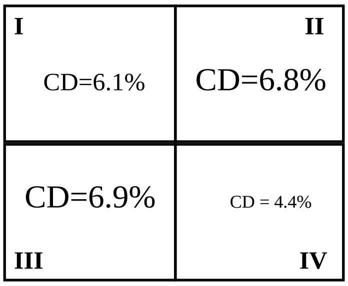

Figure 4. The average of CD of each screen quadrant (see the significant effect 'quadrant' in Table 3).

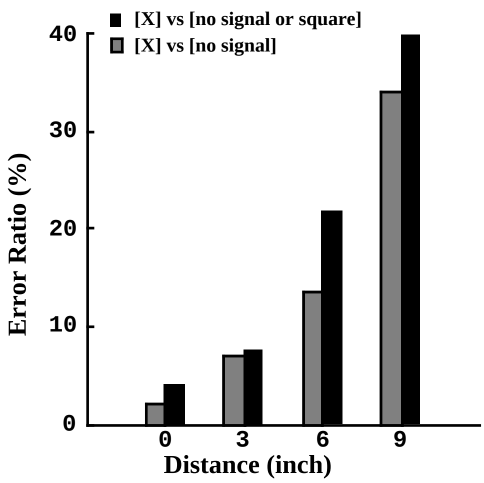

Figure 5. The error ratio (ER) plotted against the distance (D).

Overlooking an X or misinterpreting a square as an X (see Figure 5) causes the maximal detected perception errors (ER). As we can see in Figure 5 and Table 4, the error rate increases rapidly with the distance. The number of overlooked X's (cell 'b') increases between distance 0" and 3" by a factor of 3. The number of misinterpreted squares and false alarms (cell 'c') increases between distance 3" and 6" by a factor of 4.

Table 4. Contents of the cells of the 'signal detection table' and the 'error ratio' (ER). Cell 'a' and 'c' contain all cases with 'no signal' or a 'square signal'.

	a	b	c	d	ER
distance 0"	**146**	**3**	**6**	**73**	**4.1**
distance 3"	**146**	**10**	**6**	**66**	**7.5**
distance 6"	**129**	**18**	**23**	**58**	**21.9**
distance 9"	**126**	**39**	**26**	**37**	**39.8**
on average	**319**	**70**	**61**	**234**	**23.7**

Compared with XD (see Table 2) the values of ER are slightly different. The reduced sample for the analysis of variance of the dependent variable XD causes this difference.

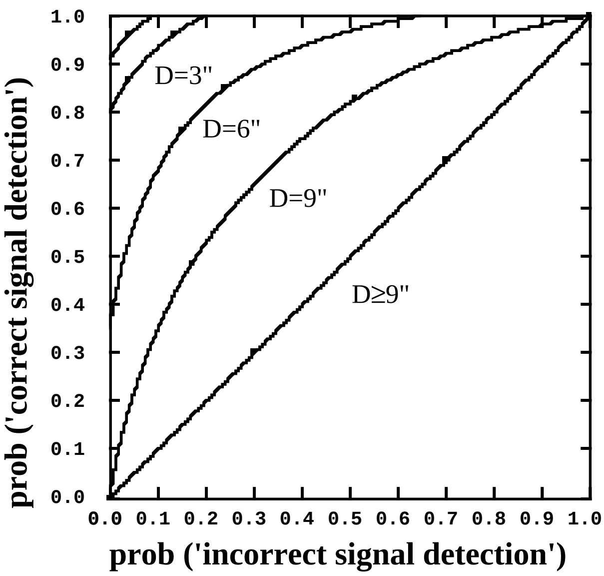

Figure 6. The 'receiver operating characteristics' (ROC) diagram regarding to Table 4. The ROC depends of the distance D.

The 'receiver operating characteristics' (ROC, see [3] and [7]) of the average user as a 'receiver' is based on the content of the 'signal detection table' (see Table 4). The probabilities prob('incorrect detection')=c/(a+c) and prob('correct detection')= d/(b+d) are calculated for each distance condition. If the 'receiver' ignores the signal and guesses, then the point describing his behaviour will fall on the diagonal running ('chance line') from the origin to the upper right-hand corner. The quality of the detection task can be estimated with a ROC diagram (see Figure 6). If different signals are easily discriminated by the 'receiver', the ROC will leave the origin with a steep slope, and it will deviate considerably from the chance line. The closer the signal is to the attention focus of the user, the better is the discriminating power.

Table 5. Contents of the cells of the 'signal detection table' and the 'error rate' (ER). Cell 'a' and 'c' contain all cases with 'no signal'; only 'false alarms' were possible.

	a	b	c	d	ER
distance 0"	76	0	3	73	2.0
distance 3"	76	10	0	66	7.0
distance 6"	76	18	0	58	13.4
distance 9"	76	39	0	37	34.5
on average	76	70	0	234	22.6

Most perceptual errors are false alarms (see Figure 5 and Table 5). If the signal appears 6" far away from the attention focus of the user, then the error ratio of 'false alarms' is 13.4% (see Table 5). The error ratio for 'false alarms' and 'misinterpretation' is 21.9% (see Table 4). The difference between these both error ratios is 8.5% and is caused by 'misinterpretations'. To reduce the perceptual uncertainty, the user moves his eyes ('focus of attention') to the screen place, where the appearance of the signal was assumed. What we now need, is a good indicator, which gives the interactive program the information, where the primary attention focus is actually on the screen. If we could find such an indicator, then we are able to present the actual message, feedback, etc. left above of the actual focus on the screen.

3 Eye Recording Experiment

In this eye movement recording experiment we have proofed the hypothesis that eye movements correlate with mouse cursor movements. To do this we carried out an eye movement recording experiment. The main assumption is that the visual focus measured with eye movement recording is an indicator for the primary attention focus of the user [5] [6].

3.1 Subjects

A total of six subjects (N=6) participated in this experiment: 2 women and 4 men with the average amount of 2175 hr's ± 1742 hr's of experience with mouse control operating a GUI. Five subjects were students of computer science at the ETH. One subject studied psychology.

3.2 Experimental Setting

To overcome the low accuracy using the NAC Eye Mark Recorder IV we projected the PC output with an overhead display on a screen of size 1.1 m x 1.4 m (see Figure 6). The subject with the eye mark recorder sat 1.5 m away from the projection screen. Using this special setting we could reduce the size of the visual angle and by that we increase the solution accuracy of the NAC Eye Mark Recorder. The subject was looking at the projection screen during the whole task solving period; to avoid a decalibration of the NAC camera the subject was instructed not to look down to the mouse device.

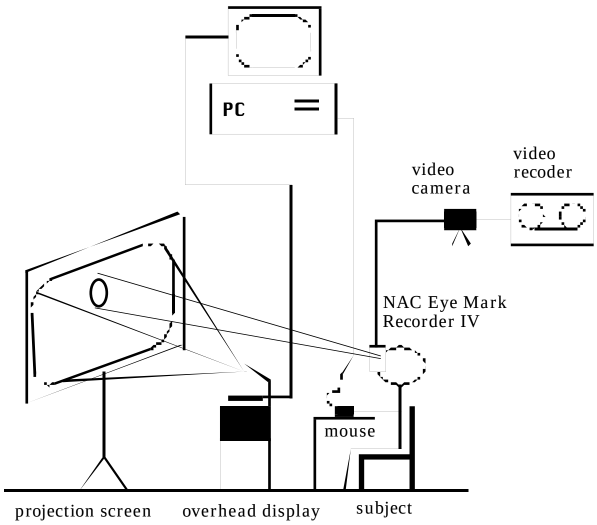

Figure 7. The experimental setting.

3.3 Tasks

Subjects were instructed to solve three tasks using the mouse:

(1) computer game: Reversi (N=4) or Solitaire (N=2);
(2) text processing (N=6): formatting a given piece of text with the text processing system Write (Windows 3.0); the subject had to read the text to get the right orientation, to select text and appropriate menu options, to press dialogue buttons and to scroll along the 70 lines of the text document;
(3) hypertext navigation (N=6): the subject had to answer five questions by navigating and searching in the help system of the Microsoft Work program.

Analysing subjects' behaviour solving the three tasks we identified the following sub-tasks:

(1) HYPER read in hypertext; 'long' click on words or other symbols (≥ 1 s).
(2) ICON click on icon on the desktop or on button in a dialogue box; click on window frame or ruler.

(3) MENU open a pull down menu, select a menu option and click.
(4) POINT point to dialogue object; double click on icon; set a Reversi stone.
(5) SCROLL scroll in a text window with button or slider.
(6) SEARCH search and look around; all other not classified sub tasks.
(7) SELECT select a piece of text.
(8) SOLI play the game Solitaire.
(9) TEXT normal text reading (≥ 1 s).

3.4 Procedure

A one-factorial test design was used. Factor A was the sub task type (HYPER, ICON, MENU, POINT, SCROLL, SEARCH, SELECT, SOLI, TEXT). All subjects had to solve the same three tasks in the same fix order (1. computer game, 2. text processing, 3. hypertext navigation). 135 min. overall task solving time was recorded. 75 min. of this material is appropriate for an evaluation, an average of 12.5 min. per subject.

3.5 Material

The experiment was run on an IBM compatible PC (Olivetti M386) with a colour screen (VGA, 14"). The standard Windows 3.0 environment was used with the delivered programs Write and Works.

3.6 Measures

To calculate the correlation of the eye movements with the mouse cursor movements the distance (d) between the visual focus and the mouse position on the screen must be measured. The NAC Eye Mark Recorder marks the visual focus as a hooklet on the video. We defined five circular regions with the following radius: 1R = 20, 2R = 40, 3R = 60, 4R = 80, 5R = 100, 6R = 120, 7R = 140, 8R ≥ 160 screen pixels. A radius of 1R is equivalent to 50 mm on the projection screen, or 9 mm on the VGA monitor. We counted with time increments of 1 s the frequencies of the mouse cursor in any of these regions (4521 data points over all subjects).

3.7 Results and Discussion

The evaluated time of each sub task and the percentage of the total evaluated recording time will be presented first. As we can see in Table 3, the sub tasks HYPER (28.2%), POINT (18.0%) and TEXT (17.5%) are the sub tasks with the largest portion of evaluated task solving time. The number of data points per sub task is equivalent to the number of seconds of the recorded part.

The content of each sub task requires more or less mouse control. Since the subjects could not use a keyboard for task solving, we can only distinguish between (1) mouse operations with visual control, (2) mouse operations without visual control, and (3) visual scanning behaviour without any mouse operations. The following sub tasks incorporate reading and looking around on the projection screen with minor mouse operations: HYPER, SEARCH, and TEXT. These three sub tasks characterise the task solving process during 48.2% of evaluated time. The other sub tasks contain intensive mouse operations: ICON, MENU, POINT, SCROLL, SOLI, and SELECT. These sub tasks are used during 51.8% of evaluated time.

For the sub tasks without mouse operations we found only a low interdependency between eye scanning behaviour and mouse movements: between 24.7% and 69.4% of the data points show the mouse cursor in the region d ≤ 4R (see Table 4).

On the other hand, for the sub tasks with mouse intensive operations we found a close correlation between the visual focus and the mouse position: 48.6% - 95.6% of the data points showed the mouse cursor in the region d ≤ 4R (see Table 4).

Table 3. Absolute and relative portions of the sub task's time of total time.

sub task	eval. time [min. : s]	percentage [%]	number of data points
HYPER	21:13	28.2	1273
SEARCH	1:52	2.5	112
TEXT	13:12	17.5	792
ICON	2:22	3.1	142
MENU	7:28	9.9	448
POINT	13:30	18.0	810
SCROLL	5:13	6.9	313
SELECT	4:09	5.5	249
SOLI	6:22	8.4	382
total	75:21	100.0	4521

Table 4. The distance between visual focus and mouse cursor position.

sub task	distance [R] (mean ± SD)	% with d ≤ 2R	% with d ≤ 4R
HYPER	4.7 ± 4.3	42.9	69.4
SEARCH	7.6 ± 5.1	21.4	37.5
TEXT	8.5 ± 4.5	15.7	24.7
ICON	2.3 ± 2.4	73.2	92.2
MENU	2.1 ± 1.4	78.6	94.2
POINT	3.0 ± 2.0	54.8	81.9
SCROLL	6.6 ± 5.1	41.6	48.6
SELECT	2.1 ± 1.3	76.3	95.6
SOLI	3.8 ± 3.1	42.4	76.2
total	4.7 ± 4.3	44.7	67.8

The sub task SCROLL seems to be an exception because only in 48.6% of the data points lay the mouse cursor nearby the visual focus. If we assume that users observe the content of the text window during 'scrolling' to control the success of the scrolling activity, then we can explain this low dependency; they do not primarily match the slider bar.

4 General Discussion and Conclusion

We can conclude from the results of the 'signal detection experiment' that the position of visual feedback on the screen must be very closed to the primary attention focus of the user. A distance over 3" should be avoided. On large screens (>14") the designer must solve the problem of estimating the actual position of the user's attention focus to avoid many unnecessary eye movements.

The results of the 'signal detection experiment' lead directly to a general design principle: the position of visual feedback must be as close as possible to the primary attention focus of the user. If the interface designer follows this principle, then he can minimise the visual scanning behaviour. The results of the 'eye recording experiment' indicate, that the mouse position is a reliable and valid indicator for the visual attention focus on the screen during mouse operations or mouse intensive tasks. We found that eye movements correlate with mouse cursor movements. If the user works with the keyboard, then we can assume, that the attention focus of the user is nearby the input cursor on the screen. This is probably true for expert users with blind typing ability. If there is no input activity, then the user reads, thinks, etc. For most of these cases the user is in an eye scanning mode, so that feedback in the centre of the screen or in reserved regions gives the user the necessary information without the feeling of being disturbed.

References

[1] Burroughs Corporation: InterPro (TM) user interface standards, Version E (March 20, 1986).

[2] DIN 66 234, Teil 3: Bildschirmarbeitsplätze – Gruppierung und Formatierung von Daten. Beuth Verlag GmbH, Berlin 30, 1981.

[3] J. P. Egan, F. R. Clarke: Psychophysics and Signal Detection. in: Experimental Methods and Instrumentation in Psychology, J. B. Sidowski , Ed., New York: McGraw Hill, 1966, pp. 211-246.

[4] P. Foley, N. Moray: Sensation, perception, and systems design. in: Handbook of Human Factors, G. Salvendy, ed., New York: John Wiley, 1987, 45-71.

[5] R. Groner, P. Fraisse, Eds., Cognition and Eye Movements. Amsterdam: North-Holland, 1980.

[6] R. Groner, G. d'Ydewalle, R. Parham (Eds.): From Eye to Mind: Information Acquisition in Perception, Search, and Reading. Amsterdam: North-Holland, 1990.

[7] W. Lee: Decision Theory and Human Behavior. New York: Wiley & Sons, 1971.

[8] OSF/Motif Style Guide: Open Software Foundation. Englewood Cliffs: Prentice Hall, Revision 1.1, 1991, p. 4/44.

[9] M. J. Staufer: Piktogramme für Computer. New York: de Gruyter, 1987.

[10] R. Teitelbaum, R. Granda: The effects of positional constancy on searching menus for information. in: Human Factors in Computing Systems, Proceedings of the CHI'83 Conference, A. Janda, Ed., New York: ACM, 1984, pp. 150-153.

[11] T. Tullis: The formatting of alphanumeric displays: a review. *Human Factors* 25, 12 (December 1981), 657-682.