The Effects of Eye Movements on Visual Inspection Performance

Mohammad T. Khasawneh[1], Sittichai Kaewkuekool[1], Shannon R. Bowling[1], Rahul Desai[1], Xiaochun Jiang[2], Andrew T. Duchowski[3], and Anand K. Gramopadhye*[1]

[1]Department of Industrial Engineering
[3]Department of Computer Science
Clemson University, Clemson, South Carolina 29634-0920

[2]Department of Industrial and Systems Engineering
North Carolina A & T University, Greensboro, NC 27411

Abstract

Quality is a key factor in business success, growth, and competitive position. One of the essential factors in quality control is the inspection task, particularly the search portion. Eye-tracking technology is a very useful tool in capturing information about the eye movements in relation to visual inspection parameters. To investigate the effect of eye movements on inspection performance, this study asked eight subjects to search for a target character on screens generated by a computer program. During this inspection task, information about the subjects' eye movements was collected to study the effect of the area covered, the number of fixations, the number of fixation points, and the filter type on the inspection performance. The results showed that the area covered during inspection did not affect the overall performance, but, on the other hand, the filter type had a significant impact.

Keywords

Visual search; eye tracking; inspection performance

1. Introduction

The manufacturing of critical application often requires a low or zero tolerance level of defects. For this reason many industries require a tedious inspection of products produced after manufacturing. This has caused much emphasis to be placed on understanding the inspection process to ensure items accepted are of requisite quality to be delivered to the customer. To this end, many studies have been conducted to see what factors may affect the performance of inspectors. By understanding these factors, interventions may be introduced to raise the quality of inspection performance. These studies have mainly focused on analyzing empirical data such as inspection time, inspection performance, training scenarios, and various other factors. All of these studies have contributed to the knowledge and understanding of the inspection process, however very little has been done to understand what the inspector is doing during the inspection process.

For some time, the ability to track the movement of the eye has existed. Crude attempts have been made as early as the 1960's using intrusive techniques such as a sclera coil and sensors placed about the eye. However, a relatively modern technique allows researchers to monitor eye movements by monitoring reflections from Infra-Red (IR) light sources and processing the data with microprocessor. The information gathered from the eye-tracking camera can then be analyzed to determine exactly where subjects are looking while performing experiments. The current eye-tracking technology offers excellent potential for monitoring inspector's eye movement during an inspection task. The eye movement data can then be analyzed in order to determine the activity of the inspector during the inspection process. From this, associations between eye movements and performance measures can be generated to develop interventions that may increase inspection performance. Also, the recording of eye movements can provide useful information on whether sufficient time is being allowed for the inspection of a product and, in particular, whether a product is being given adequate visual coverage [6]. Therefore, the objectives of this study are to investigate the following hypotheses and research questions that are related to inspection performance and eye movement data: 1) is there a difference between the area covered, number of fixations, and number of fixation points between a 2-tap and 5-tap filter; 2) is there a difference in screen area covered for subjects with different inspection accuracies; and 3) is

* Corresponding author, Tel: 864 656 5540, e-mail: agramop@ces.clemson.edu.

there a difference between the area covered, number of fixations, and number of fixation points between screens with misses and screens with hits for a 2-tap and 5-tap filter.

2. Methodology

2.1 Subjects

In this study, a selected group of 8 students was chosen at random from the population of Clemson University. After signing a consent form, subjects were asked to carry out the visual inspection task in a simulated experiment. Using student subjects is an assumption that can be validated according to Gallway and Drury [3], where it has been proven there are minimum differences between real world inspectors and students in simulated tasks.

2.2 Stimulus Material & Inspection Task

The inspection task in this experiment was a simulated defect search task in which subjects were asked to search for a defect on screens generated by a computer program written in Visual Basic 6.0. The screens were randomly generated using a set of ASCII characters (W, N, M, A, X, Y, K, Z) as the background with a density of 20%, and the subjects' task was to search each screen for a possible defect, V, which is the target character. The viewing screen is a 27-inch HDTV with a screen resolution of 600 × 450. The experiment was conducted on a Silicon Graphics Dual Rack Onyx2 Infinite Reality system. The system has the following specifications: 8 R10000 Mips Processors with 4 Mb onboard cache, 1 Gb Main Memory, 45 Gb Disk Storage, 2 Graphics pipes (1 per rack), 4 Raster Managers (RM) per Pipe, 64 Mb Texture Memory per RM, 2-25 inch HDTV Monitors per Pipe.

The subjects performed the inspection task at a 30-inch distance from the monitor. The inspection was a machine-paced task, which only involved the visual search component. During visual search, screens were presented to subjects and their task was to locate the defect (i.e. target character). When they detect a defect, the subjects were asked to stop searching and fixate on the defect until the next screen appeared. Each inspection task consisted of 10 randomly ordered screens, of which 80% contained the target character. A sample screen from the stimulus material is shown in Figure 1 below.

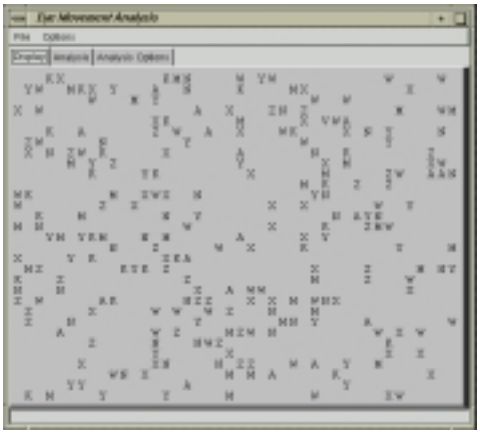

Figure 1: A sample screen of the stimulus material.

A paced task is one in which a time limit has been imposed, while schemes of pacing deal more with the degree of control one has over the task. Three type of pacing have been discussed in the literature: machine paced, self paced, and unpaced. Since some of these terms have been used interchangeably, often what has been presented in the literature is not consistent [7]. A machine-paced task is defined as a fixed time in which a defect may be detected. The same amount of time is allocated whether or not a defect is found or not. Self-paced is when a maximum time limit is set, although the inspector may choose to go on before the time limit is reached. Finally, unpaced inspection occurs when there is no time restriction place on the inspector. Therefore, for the purpose of this study a machine paced inspection task was simulated.

2.3 Pilot Study

A pilot study was conducted in order to set appropriate pacing times. Three subjects were chosen at random and given a simulated inspection task. Each subject was told to inspect as quickly and accurately as possible. The

accuracy levels, search times, stop times, and inspection time were recorded. The average inspection time of each subject was 15 seconds per screen. This time was used as the pacing time for the actual experiment.

2.4 Experimental Design

The study used a single factor within-subject design. Table 1 shows the layout of the design. The sequence of the screens presented to each subject was randomized to cancel any order effects that may have occurred. The experimental design was a within subjects design wherein all the subjects underwent same set of experimental conditions.

Table 1. Experimental Design (A, B, C, D, E, F, G, H, I, and J correspond to the screen number)

		Subjects							
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	1	A	B	C	D	E	F	G	H
	2	B	C	D	E	F	G	H	I
Trial	**3**	C	D	E	F	G	H	I	J
	4	D	E	F	G	H	I	J	A
	5	E	F	G	H	I	J	A	B
	6	F	G	H	I	J	A	B	C
	7	G	H	I	J	A	B	C	D
	8	H	I	J	A	B	C	D	E
	9	I	J	A	B	C	D	E	F
	10	J	A	B	C	D	E	F	G

2.5 Procedure

On the experiment day, each subject was required to sign a consent form and complete a demographic questionnaire. Following this step, instructions were read to the subjects to ensure their understanding of the experiment. On completion of the study, subjects were debriefed and thanked for their participation.

3. Results

3.1 Filter Comparison

Because the factors investigated relay heavily on what type of algorithm/filter is used for calculation, a comparison is made to determine if different algorithms/filters differ significantly. The comparison is made between a 2-tap and 5-tap filter, which determines the percent area covered, number of fixations, and number of fixation points (see Table 3). Result from a means comparison test of two independent samples showed that all three factors differ significantly depending on whether a 2-tap or 5-tap filter was used (see Table 4).

Table 3 (Statistical Data for Mean Comparisons)

	TAP	**N**	**Mean**	**S.D.**	**SEM**
PERCENT	2	21	23.06	14.97	3.27
	5	75	12.77	7.99	0.92
NOSFIX	2	21	209.52	120.44	26.28
	5	75	18.13	11.80	1.36
NOSFIXPT	2	21	212.57	120.48	26.29
	5	75	438.89	232.06	26.80

Table 4 (Independent Samples Test)

	T	**df**	**Sig. (2-tailed)**
PERCENT	4.210	94.000	0.0000585
NOSFIX	7.272	20.108	0.0000005
NOSFIXPT	6.029	64.363	0.0000001

3.2 Area Covered Comparisons based on Inspection Accuracy

To determine if there is a difference in the area covered based on inspector accuracy, an analysis of variance was conducted on the area covered between group accuracy, which is defined as the number of defects found. The

results compared the area covered between inspector that found two, three, and four defects in each experiment. The comparison was made for both a 2-tap and 5-tap filter. An ANOVA showed that both the 5-tap filter and 2-tap filter comparison was not significant at the p-value 0.05 (see Tables 5-6). However it is interesting to note that a LSD post hoc analysis for the 5-tap filter comparison showed a significant difference between the two-defect group accuracy level and four-defect group accuracy level (p=0.042). Table 7 shows the mean and standard deviation for the percentage of area covered for each corresponding accuracy. It can be seen that the maximum number of defects found was four and the minimum was two.

Table 5 (5-Tap Filter)

	Sum of Squares	**df**	**Mean Square**	**F**	**Sig.**
Between Groups	364.17	2	182.09	2.94	0.06
Within Groups	4525.15	73	61.99		
Total	4889.32	75			

Table 6 (2-Tap Filter)

	Sum of Squares	**df**	**Mean Square**	**F**	**Sig.**
Between Groups	209.10	2	104.55	0.34	0.71
Within Groups	22946.82	75	305.96		
Total	23155.92	77			

Table 7 Descriptive statistics for both filter types used.

Accuracy 5-tap	**N**	**Mean**	**S.D.**		**Accuracy 2-tap**	**N**	**Mean**	**S.D.**
2	46	14.30	8.64		2	48	29.26	18.32
3	20	10.69	6.97		3	20	28.56	16.69
4	10	8.63	5.22		4	10	33.87	14.46
Total	76	12.61	8.07		Total	78	29.67	17.34

3.3 Missed Defect Comparison

Screens in which the inspection fixated on the target character but did not identify a defect are classified as a miss, inspectors that fixated on a defect and identified a defect are classified as a hit, and screens in which the inspector did not fixate on the target and did not find a defect are classified as a no-miss. A comparison between hits and misses is made for percent area covered, number of fixation points, and number of fixations. For the 2-tap comparison, a significant difference was found for all three factors (see Tables 8-9).

Table 8 (Statistical Data for Missed Defect Mean Comparisons 2-tap)

Factor	**Miss 2-tap**	**N**	**Mean**	**S.D.**
% Area Covered	No	53	28.30	15.82
	Yes	7	42.35	15.74
# of Fixations	No	53	238.09	121.85
	Yes	7	345.71	118.01
# of Fixations Points	No	53	242.21	122.12
	Yes	7	350.14	118.77

Table 9 (Independent Samples Test for Missed Defect 2-tap)

Factor	**Sig.**	**df**	**t**	**Sig. (2-tailed)**
% Area Covered	0.983	58	2.209	0.031
# of Fixations	0.578	58	2.203	0.032
# of Fixations Points	0.594	58	2.204	0.032

For the 5-tap comparison, no significant difference was found for the three factors (see Tables 10-11)

Table 10 (Statistical Data for Missed Defect Mean Comparisons 5-tap)

Factor	**Miss 5-tap**	**N**	**Mean**	**S.D.**
% Area Covered	No	36	13.11	8.74
	Yes	24	10.82	5.46
# of Fixations	No	36	35.03	102.71
	Yes	24	15.75	8.89
# of Fixations Points	No	36	419.25	217.95
	Yes	24	447.00	273.56

Table 11 (Independent Samples Test for Missed Defect 5-tap)

Factor	**Sig.**	**df**	**t**	**Sig. (2-tailed)**
% Area Covered	0.03	58	1.14	0.26
# of Fixations	0.16	58	0.91	0.36
# of Fixations Points	0.09	58	0.44	0.66

Figure 2 below shows a sample screen when the subject was able to find the target character (V). Also, Figure 3 shows a "miss" where the target character was not detected although the subject's scan path passed through the region of interest. Figures 4 shows the effect of using a 5-tap filter for analyzing the data, where in the defect was not detected. On the other hand, figure 5 shows the same sample screen where the target character was detected when a 2-tap filter was used.

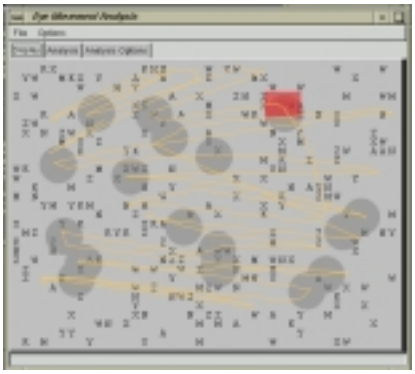

Figure 2: A sample screen that shows a defect that was found.

Figure 4: A sample screen that shows a defect that was not found using a 5-tap filter.

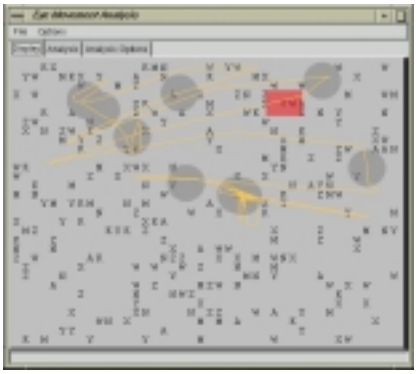

Figure 3: A sample screen that shows a defect that was missed.

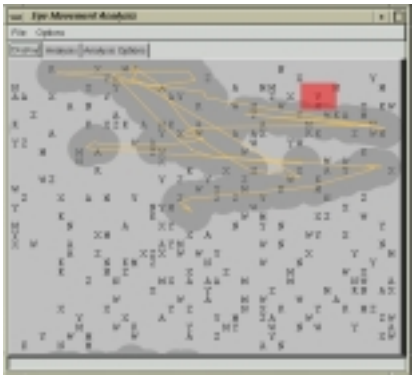

Figure 5: A sample screen that shows a defect that was found using a 2-tap filter.

4. Discussion and Conclusions

This study investigated the effect of eye movements on inspection performance. The result from the study showed a significant difference in the percent area covered, number of fixations, and number of fixation points between a 2-

tap and 5-tap filter. This indicates that the type of filter used for analysis will affect the results. Therefore careful consideration should be taken when determination the type of filter that will be used for analysis. From Models of visual search [4,5] and search data [2], a major factor affecting search performance is the area is searched, or the number of non-target elements in the field. To investigate this fact, and based on the data obtained, the results showed that no significant difference in percent area covered was found for subjects with different inspection accuracies. This indicates that inspection performance does not correlate with percent area covered. Therefore "better" inspectors do not necessarily cover more or less area when searching for a defect. This might be due to the fact that some subjects might have performed a systematic search where the location of the target would be relevant. That is, if the target is close to the starting point of the subject's scan path, the area covered would be minimum. However the memory less property of the random search made the actual percentage of area covered less due to the overlap regions. There was also no significant difference in the number of fixations and number of fixation points based on inspection accuracy for either the 2-tap or 5-tap filter. This indicates that "better" inspector do not fixate more or less than "worse" inspectors.

Analysis of the data for a 2-tap filter showed a significant difference between the areas covered, number of fixations, and number of fixation points between screens where there was a missed defect and screens where no miss occurred. However, analysis of the 5-tap filter data did not show any significant difference between the miss and no-miss screens. This fact once again emphasizes the filter dependency of the results because a 5-tap filter reduces variation of data more than a 2-tap filter. Therefore a significant difference must exist before 5-tap filter analysis will yield different results. Furthermore, table 8 shows that subjects with missed defects covered more area than those with no misses. This is due to the fact that subjects have more time to reexamine the screen and hence cover more area.

Machine paced inspection in industry offer certain economic advantages such as the minimization of work in progress, maximization of floor space usage, and simplification of the organization of supplying components to the right place at the right time [1]. However, because under this type of paced condition, operators are required to complete each task within a rigidly fixed time, certain ergonomics principles are lacking (i.e. stress originates from forcing longer than standard work cycle times into a rigidly fixed cycle time). This fact may have affected the subject's performance during the inspection task, specifically the subjects may have fixated on the target but due to time stress may not have detected the presence of the target. Based of the results obtained from the experiment we can conclude the following:

- Filter type has a significant effect on the result of the study.
- The area covered during inspection was not affected by the subject's performance.
- Screen with misses and those with no-misses did differ significantly in terms of area covered, number of fixation, and number of fixation points for a 2-tap filter.
- Screen with misses and those with no-misses did not differ significantly in terms of area covered, number of fixation, and number of fixation points for a 5-tap filter.

There are several other extensions that can be explored to improve the current study. First, the effect of different pacing schemes on the eye-movement parameters can be investigated. Secondly, more research needs to be conducted in order to determine what type of filter best applies for eye-movement analysis. Thirdly, other performance measures such as mean inspection time, mean stopping time, mean search time, and mean inter-fixation time should be investigated in order to obtain a comprehensive view of their effect on visual search strategies.

References

1. Belbin, R. M. and Stammers, D., "Pacing stress, human adaptation and training in car production," Applied Ergonomics, 1972.
2. Drury, C. G. and Clement, M. R., "The effect of area density, and number of background characters on visual search, Human Factors, 20, pp. 597-602, 1978.
3. Gallwey, T.J. and Drury, C.G., "Task complexity in visual inspection," Human Factors, 1986.
4. Karwan, M., Morawski, T. B., and Drury, C. G., 'Optimum speed of visual inspection using a systematic search strategy," IIE transaction, 1995.
5. Morawski T., Drury, C. G., and Karwan, M., "The optimum speed of visual inspection using a random search strategy," IIE transaction, 1992.
6. Megaw, E. D., and Richardson, J., "Eye movements and industrial inspection," Applied Ergonomics, 1979.
7. Sandra K. Garrett, Brian J. Melloy, and Anand K. Gramopadhye, "The effects of per-lot and per-item pacing on inspection performance," International Journal of Industrial Ergonomics, 2001.