Evaluation of Perception Performance in Neck Dissection Planning using Eye Tracking and Attention Landscapes

Oliver Burgert [a*], Veronika Örn [a], Boris M. Velichkovsky [b], Michael Gessat [a], Markus Joos [b,c], Gero Strauß [d], Christian Tietjen[e], Bernhard Preim[e], Ilka Hertel[d]

[a] BMBF-Innovation Center Computer Assisted Surgery ICCAS, University of Leipzig, Philipp-Rosenthal-Strasse 55, 04103 Leipzig, Germany;

[b] Institute for Psychology III, Technical University of Dresden, 01069 Dresden, Germany;

[c] IMD Interactive Minds Dresden GmbH, Oberwachwitzer Weg 13, 01326 Dresden, Germany;

[d] Dept. of ORL, University of Leipzig, Liebigstrasse 10-14, 04103 Leipzig, Germany;

[e] Faculty of Computer Science, University of Magdeburg, 39106 Magdeburg, Germany.

ABSTRACT

Neck dissection is a surgical intervention at which cervical lymph node metastases are removed. Accurate surgical planning is of high importance because wrong judgment of the situation causes severe harm for the patient. Diagnostic perception of radiological images by a surgeon is an acquired skill that can be enhanced by training and experience. To improve accuracy in detecting pathological lymph nodes by newcomers and less experienced professionals, it is essential to understand how surgical experts solve relevant visual and recognition tasks. By using eye tracking and especially the newly-developed attention landscapes visualizations, it could be determined whether visualization options, for example 3D models instead of CT data, help in increasing accuracy and speed of neck dissection planning. Thirteen ORL surgeons with different levels of expertise participated in this study. They inspected different visualizations of 3D models and original CT datasets of patients. Among others, we used scanpath analysis and attention landscapes to interpret the inspection strategies. It was possible to distinguish different patterns of visual exploratory activity. The experienced surgeons exhibited a higher concentration of attention on the limited number of areas of interest and demonstrated less saccadic eye movements indicating a better orientation.

Keywords: Attention, Eye Tracking, Image Perception, Observer Performance Evaluation, Tumor Diagnostics

1. INTRODUCTION

In neck dissection surgery, the successful removal of lymph node metastases is essential for increasing the chance of the patient's survival. In order to reduce the risk for the patient, a proper planning of the intervention has to be performed. Even more, surgical planning evaluates respectability (the feasibility of complete and safe tumor removal) and leads to the decision whether an operation is useful at all for the patient. The position of large metastases should be accurately determined, so that a potential harm to structures at risk, e.g. Arteria Carotis, can be reduced. The detection of suspicious lymph nodes in CT images is a difficult task, even for experienced radiologists and surgeons. Expert radiologists segment the lymph nodes and other anatomical structures necessary for the preoperative planning and the intraoperative orientation. 3D models of these segmentations are used afterwards for surgical planning.

Until now, it has not been convincingly shown that 3D models increase the quality of surgical planning. Evaluation of parameters such as the set-up duration in dependence of special visualization, the optical quality of images, and the surgeon's experience were essential to determine appropriate visual options in order to improve the analysis accuracy of radiological images and data. We started using additional 3D visualizations of the neck area in order to test if 3D models improve the accuracy of intervention-planning and reduce stress load for surgeons. In a previous study we used a questionnaire containing subjective measurements of the cognition performance of surgeons with different levels of experience. The results gave us guidelines for selecting visualization parameters[1]. However, for an optimization of performance, the visual perception of the surgeons inspecting different visualizations must be understood more

* oliver.burgert@iccas.de; phone +49 341 97-12000; fax +49 341 97-12009; www.iccas.de

thoroughly. In the present study we decided to use eye tracking and the so-called attention landscapes[2] software visualizing the exact distribution of observers' attention for the purpose of evaluating surgeons' perceptual performance. Using eye tracking technology, we aimed at selecting the most beneficial imaging options in order to improve the accuracy and speed of surgical planning as well as of overall perception performance. This technology reveals areas the viewer looks at, it shows how long his/her gaze dwells on them, and it illustrates the exact way of perceptual interpretation of the displayed information.

The earlier and more accurate detection of malignant lymph nodes increases the chance for a complete cure. To avoid human errors related to the introduction of new image visualizations in surgery, it is necessary to verify technical specifications and assess the system's usability in the clinical context.

2. METHODOLOGY

2.1 Eye Tracking Technology and Attention Landscapes

There is a variety of methods to use eye tracking: they depend on the research field, issues, and parameters which need to be evaluated. There are remote and head-mounted eye tracking systems. For the measurement of eye movements (the detailed sequence of fixations at a displayed image) in most remote eye tracking systems, the PCCR (Pupil Center Corneal Reflection) method is used. The point of regard can be computed by matching the PCCR vector to a set of known positions on a computer screen: it is usually performed with a calibration routine. In the case of tracking, recording, and analysing surgeons' eye movements, remote eye tracking systems are the method of choice, as they are unobtrusive, with nothing attached to the subjects and thereby not influencing normal visual behaviour of subjects[3]. Analysing the scan paths helps to find the most comfortable and effective way of visualization for greater accuracy in reading radiology images. We used a binocular remote system which consists of software for recording and analysing gaze direction in the context of the displayed image and hardware including standard PC components complemented by an optics module (two eye cameras and an IR light source) from LC Technologies, Inc., USA. The system is shown in Figure 1.

Fig. 1: Video-based 3D image evaluation using an eye tracking system (courtesy of LC Technology, Inc. and IMD GmbH).

EyeGaze, the binocular eye tracking system used in our experiments, is a remote system. It was positioned below the computer monitor and aimed at a seated subject in order to keep an optimal contact between the installed cameras, light source, and viewer's eyes. Data is lost if subjects move out of the eye tracker's range (approx. 8 degrees of arc in each direction), or look too far away from the eye tracker optics. The recording of eye movements starts after the calibration of the eye tracker. This system was used to record and analyse eye movements of surgeons, while they solved intervention planning tasks on the basis of 3D or CT neck images.

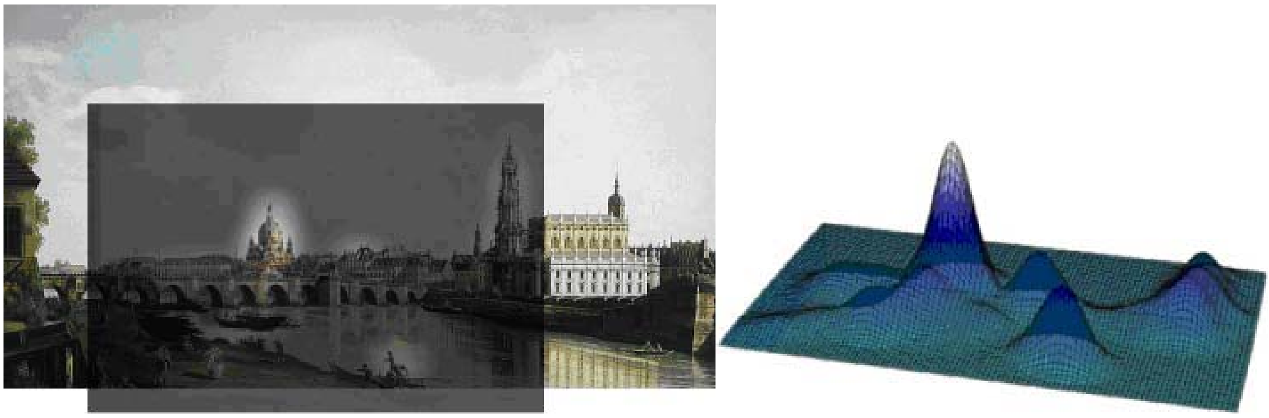

Fig. 2: An example of the original painting is in the background, an empirically measured and computed attention landscape on the right, and the visualisation is superimposed on the painting[9].

While eye tracking technology has been extensively used in medicine and elsewhere for interpreting perception of imaging data[3,4,5,6,7,8], a new aspect of this study was the application of the so-called attention landscapes in medicine. The first step towards the method was made by Velichkovsky, Pomplun, and Rieser[2] with their proposal to elucidate the perception of ambiguous paintings (Archimboldo, Dürer and Escher) by rendering of these pictures in terms of the distribution of visual fixations. The idea behind this approach is twofold. First, due to the saccadic suppression, visual processing is limited by the duration of fixations. Second, due to the different density of receptors in the retina, each fixation resolves only a narrow area of the visual scene that can be approximated by a Gaussian distribution with the standard deviation of 1 degree of arc. In this way, it could be made visible when an observer perceived, e.g. an angel or a devil in Circle Limit IV by Escher. We illustrate the approach with Canalettos' View of Dresden in Figure 2. The attention landscape analysis is a standard feature of the NYAN eye tracking data analysis suite by IMD GmbH (Dresden) used in our investigation.

As it was repeatedly stressed in the literature[4,9], knowing which physical features attract the visual attention of surgeons may help researchers in developing image analysis methods for identifying lesions and processing tools that may enhance the visibility of those features of interest.

2.2 Test Set-up

To compare the results of this study with the one evaluating the cognitive performance[1], we used the same patient data: five CT data of different patients were used to create 32 different visualizations[10]. We used real-world CT data brought by the patients from different radiological offices, and the CT data was sometimes of poor quality. From these 32 visualizations and 5 CT datasets we selected four 3D visualizations (no. 7, 9, 13, 24) and two CT datasets to assure a short but concentrated session. The visualizations differed in the coloration scheme and application of transparency schemes as illustrated in Figures 3 to 6. They were presented in frontal, left, right, and dorsal view that could be simultaneously presented on the same display. For CT data, we selected two different views (transversal and cranial), which could be presented on two levels of image quality.

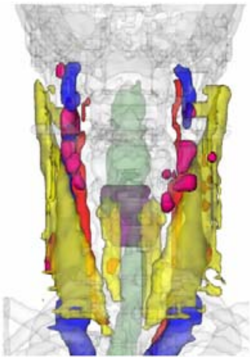

Fig. 3. Model 24, abnormally colored, transparency of muscles and bones, frontal view.

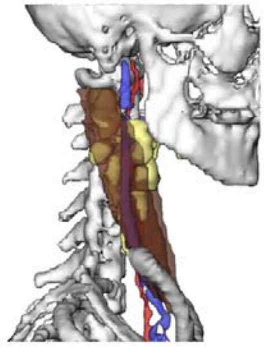

Fig. 4. Model 9, naturally colored, transparency of muscles, view of right side.

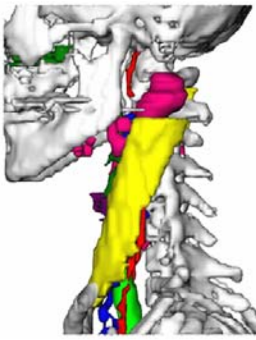

Fig. 5. Model 13, abnormally colored, view of left side.

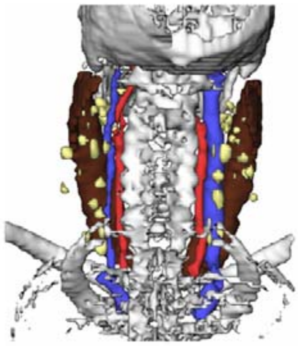

Fig. 6. Model 7, naturally colored, dorsal view

Thirteen surgeons with different levels of expertise (see Table 2) were asked to complete a questionnaire containing questions normally answered during the surgical planning phase such as the number of suspicious lymph nodes at different levels of the neck, or the infiltration of anatomical structures like the Arteria Carotis. This task had to be performed first on the 3D visualizations and then on the CT data. Subjective questions concerning the perception performance had to be answered on a 5 points response scale, from -- (completely inaccurate) to ++ (completely accurate).

Besides this questionnaire, we used the “EyeGaze Analysis System”, an eye tracking solution provided by IMD Interactive Minds GmbH, Dresden, Germany. For the processing of recorded eye movements, we used NYAN eye tracking data analysis suite Version 1.3. Among other features, we used scanpath analysis, area of interest analysis, attention landscapes, single subject and group comparison analysis. We defined four areas of interest (AOI 0 - AOI 3) in each image; each AOI covers one view (frontal, left, right, dorsal) of the model. For each AOI, the following calculations have been performed: AOI Dwell Time absolute in percent, Number of Fixations per AOI, Mean Fixation Duration per AOI in milliseconds, Time to First Fixation of AOI in milliseconds and a summarized Duration of Fixation for each image in milliseconds. The particular relevance of these indices has been demonstrated in the previous applied research based on eye tracking technology.[8]

After calibration of the eye position data and before starting the recording of eye movements, personal data and distinguishing features of the tested surgeons such as an anonymous code of four letters, the sex, and especially the level of experience have been determined. During the test set-up five questions concerning each image had to be answered: (1) number of suspicious lymph nodes with (2) their specific anatomical ‘level’, (3) the classification of the tumor extent, and (4) its possible infiltration as well as (5) the ease of recognition of different anatomical structures. After each session, further three questions about subjective opinions about 3D visualization and CT data were asked. These questions addressed the previous experience with 3D images, the subjective excess load while working with 3D images, and the possible advantage in having access to a view on both types of images, 3D and CT, at the same time. The duration of the test set-up presenting four 3D models and two CT data was not limited. In average, it took between 20 to 30 minutes for each surgeon. The presentation time for 3D images and CT data was counterbalanced. Images of both types followed each other after a short break of approximately one minute. The surgeons were able to move the CT films independently in order to increase accuracy in recognizing pathological lymph nodes in the neck region. This was done with the intention of reconstructing the similar conditions as in everyday clinical settings.

3. RESULTS

The inspection of each 3D image took mostly from 3 to 5 minutes; the whole eye tracking session lasted about 20 to 30 minutes per surgeon. Since the 3D visualizations of the patients were using a standardized screen layout (frontal, left, right, and dorsal view) (Fig. 7), it was possible to compare viewing strategies even across models. In addition, it was possible to compare the inspection strategies of the experienced surgeons with those of the novices.

Every surgeon started inspecting the model using the frontal overview (AOI 0) positioned on the top left of the display. After a rough orientation of approximately 1-3 seconds, the gaze line led to the image presenting the left side of the neck (AOI 1) on the top right and went back to the frontal view. The image showing the dorsal sight (AOI 3) (lower right) was - by the majority of the subjects - inspected as the third one before the view returned to the overview image. Nearly each surgeon ended the inspection on the right side of the neck (AOI 2) down on the left of the display between 6 and 10 seconds after the start of the recording of eye movement. The preferred view of the experts seems to follow a similar order as the one of the novices. Compared to the novices, the experienced surgeons showed a significantly lower frequency of the saccadic eye movements (changes in the surgeon's gaze direction, Fig. 7) between the four possible views of the presented model, indicating a faster overview of important information and a better spatial orientation in the images. The Mean Fixation Duration (the longest view at one point of the whole set) in each of the four Areas of Interest (Fig.8) was positively correlated with the experience of the surgeons. In Fig. 8, each coloured circle represents one fixation; the diameter is correlated to the fixation duration. The novices showed an insignificantly higher Number of Fixations per 3D model. Special patterns of eye movements were not recognizable.

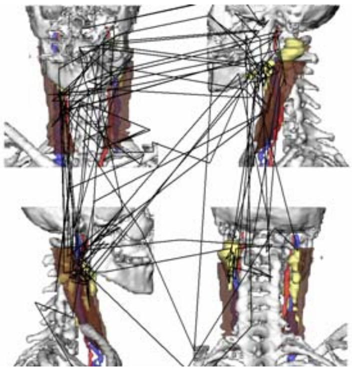

Fig. 7. Scanpath with trajectories of saccades.

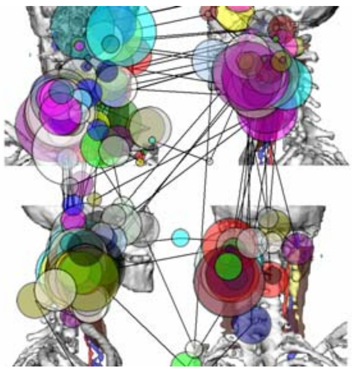

Fig. 8. Visual fixations with visualization of their duration.

Average Times to First Fixation were the shortest in the AOI 0 as the first inspected region for all four 3D models. This region also showed the longest Absolute Dwell Time for three of the four visualization models (refer to Table 1). The length of the Dwell Time duration took about 1 and 2 ½ minutes in the preferred AOI. The Number of Fixations per AOI varied between the surgeons without showing a significant correlation to the Dwell Times in the same region.

Table 1. Dwell Times for each area of interest in the 3D model 24 for one of the surgeons in milliseconds.

Surgeon	3D model	Area of Interest AOI	Dwell Time in ms
DSDM	24	0	52336
		1	13081
		2	11431
		3	9319

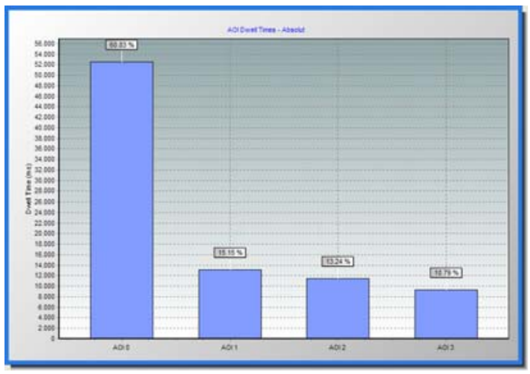

Table 2. Time to First Fixation for each area of interest of the different 3D models for the ORL surgeons.

Surgeons' Code		AAHT	AAHT	BAAD	GAUH	GDHT	CAKS	HTCA	HAET	MNED	MSJG	DSDM	JARR	SEUH
Expert		yes	yes	yes	yes	yes	Yes	yes	yes	no	no	no	no	no
Sex		m	m	f	f	f	F	m	m	f	f	m	f	f
AOI	Model													
0	7	500	742	467	500	225	2800	400	333	83	1583	200	550	333
1	7	192	1625	3200	192	43200	2150	11752	1882	5000	1050	458	53467	1466
2	7	3101	3133	7350	3101	12843	77701	1583	2732	3350	5649	4991	13300	3233
3	7	21225	7324	8384	21225	43350	417	48071	4365	3700	4565	10633	52834	25196
0	9	584	366	1350	584	358	3767	1666	1183	900	1200	442	1066	383
1	9	1676	1982	3167	1676	4556	1067	2183	2549	333	1900	9751	783	1150
2	9	2342	233	6935	2342	27191	9099	433	1916	8082	5867	709	11565	2700
3	9	2084	8850	5651	2084	22716	8066	6282	16065	4950	2000	3419	58032	4283
0	13	200	50	400	200	308	2451	2500	583	633	817	225	1167	167
1	13	1066	1342	117	1066	1807	367	6234	1332	2950	1467	542	3366	1517
2	13	1916	1776	2683	1916	12564	34638	1434	1649	5400	5684	3865	13049	5217
3	13	1599	2184	3083	1599	7016	3919	12017	5981	4600	3084	4165	20999	5067
0	24	808	58	133	808	358	100	667	683	133	1483	642	750	584
1	24	425	1617	1949	425	21911	3250	2467	1500	1700	800	1217	99948	2017
2	24	3634	1809	15648	3634	22527	42782	1400	1950	3300	4316	225	8483	6485
3	24	3176	2709	2349	3176	25421	2933	46017	17217	2450	3216	8592	9600	18853

Thus, the analysis of both scanpaths for saccadic eye movements as well as the parameter of the number of fixations led to the conclusion that more experienced surgeons needed less orientation between the images. This conclusion, however, says nothing about the form of attention distribution. Therefore, after the registration of eye movements, we applied the attention landscapes software that allows to process images in terms of distribution of visual attention as manifested in the position and duration of fixations. An example of such visualization for one of the images and an individual expert surgeon is shown in Figure 9.

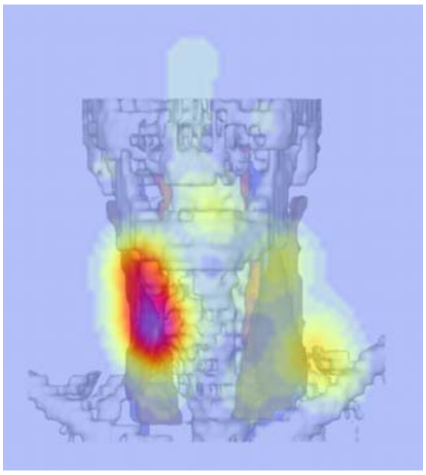

Fig. 9. Visualization on the basis of attention landscapes: The focus of attention in the case of an expert surgeon evaluating a 3D image is shown.

The feature Attention Landscapes allows to evaluate the main attention of every surgeon during the set crossover the defined Areas of Interest and visualizes it simultaneously. Comparing the recorded attention between the expert surgeons and the less experienced professionals in defined areas of interests shows a highly significant difference testifying to a stronger concentration of attentional recourses on the local areas of images in the case of expert surgeons (see Fig. 10 and 11).

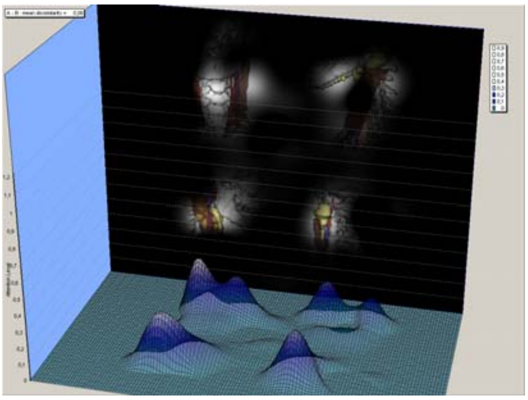

Fig. 10. Visualization of differences in the distribution of attention: Novices minus experts.

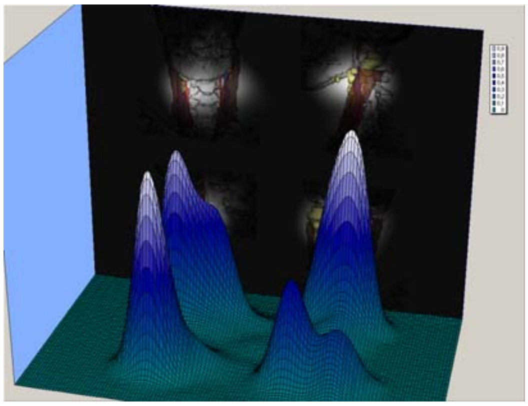

Fig. 11. Visualization of differences in the distribution of attention: Experts minus novices.

A special problem was related to the evaluation of CT data. With the intention of keeping the simulated inspection of a CT film as close as possible to the clinical conditions, the surgeons were allowed to move the CT data. As a result, the evaluation of the available features was not possible because the "scroll activities" during the set-up could not be recreated. The current software does not allow to classify scanpaths, saccades, and fixations to a special anatomic level of the neck, and the interactive handling of the CT data prohibits to define single areas of interest. The content of the screen changes too fast for a proper evaluation. In the future we will use more refined techniques to analyse the gaze in the second part of the evaluation.

The inspection duration of CT data with the goal of recognizing defined lymph node levels in neck region as well as distinguishing pathological lymph nodes from background tissue and other anatomical structures like vessels depends significantly on the surgeon's expertise. Replicating earlier reports, this fact indicates that expertise is likely to be a combination of specific visual and cognitive skills derived from medical training and experience in detecting and determining the diagnostic importance of radiographic findings[6,7].

So far, we have analyzed the subjective part of the questionnaire for CT inspection. Both groups rated their own perception performance as being better using 3D visualization. However, nearly all of the surgeons would prefer to have a view of both visualizations–3D and CT–simultaneously. 60 percent of all surgeons state that there is no excessive cognitive load while inspecting 3D images compared with inspecting CT data. The analysis of the questionnaires revealed that the amount of detected lymph nodes and the recognition of infiltration in the individual images varied considerably independent from the surgeons' expertise. This could be a result of the unusual and ambiguous visualization of lymph nodes and anatomical structures, especially in those 3D models presenting transparency of muscles and bones. The difficulty of recognizing anatomic levels, detecting lymph nodes, and distinguishing the infiltration in the particular CT data and 3D images (change in coloration and transparency) was assessed by both, experienced and non-experienced surgeons as relatively similar. Due to the eye movement data it was possible to distinguish different exploration patterns (the frequency of scanpath saccades), but no special patterns of result (e.g. Dwell Times absolute per AOI) for the experienced or inexperienced surgeons were recognizable.

4. CONCLUSIONS

The results from the present study are useful for developing software with new analysis features and improving the already existing eye tracking analysis applications like e.g. the evaluation of movable CT data. By transfer of currently available features of eye tracking technology to other analysis tasks, more details about patterns of looking at diagnostic images should be investigated. Together with the developers of this system, we define cases for further studies and the requirements for analysis software. These iterations lead towards the development of more flexible software and the use of more refined techniques.

Eye trackers could be used for optimising visualizations and screen layouts as well as for the training process for surgeons without expertise who need comparable examples of image perception and recognition tasks of professionals. Thus, the recorded scanpaths of experts can be used for teaching novice surgeons the best exploration strategies in a very intuitive manner[6,7]. At the same time, improving the accuracy of detecting pathological lymph nodes and reducing stress load, a simultaneous presentation of both image types-CT and 3D-could be an advantage, especially for experienced surgeons. Compared to novices, the additional 3D visualization supports their anyway faster and more accurate detecting of suspicious lymph nodes resulting from more experience with CT data inspection[11]. The analysis of inspection strategies of surgeons exploring 3D visualizations leads towards the development of better visualization strategies, an improved design of interfaces, and the generation of new models. This includes screen layout and selection of visualization parameters for the models. In this study, nearly every surgeon started inspecting each model using the frontal overview (AOI 0) positioned on the top left of the display. In a later study, with a changed order of the four possible views of all used 3D models it could be evaluated if the reason for this might be the usual left-to-right reading pattern or the high diagnostic value of the frontal view.

For an optimal analysis, it is important to record perception, cognition, and emotional evaluation of the surgeon simultaneously. For instance, in the setting we have chosen for the present study, it is not clear whether the surgeon looked very shortly at a certain area because he immediately saw a relevant structure, or whether he just did not notice it. In this and other tasks, further evaluation should follow. One such further development is related to the recent differentiation of focal and ambient attention modes as manifested in attention landscapes[9,12]. In a still other perspective, an application of research methodology from affective neuroscience and neuroeconomics is conceivable. This methodology allows an objective evaluation of emotional attitudes of perceivers toward different aspects of the visual image[13,14]. By combining these methods, we could come closer to elucidation of the psycho-physiological basis for diagnostic decisions in order to efficiently support them with appropriate technical tools and organizational procedures.

ACKNOWLEDGEMENTS

The authors wish to thank all the collaborators at the Innovation Center Computer Assisted Surgery (ICCAS), Medical Faculty, University of Leipzig, and at the Unit of Applied Cognitive Research, Technical University Dresden. The investigation described in this paper was partly supported by the EU Commission NEST-Pathfinder Project PERCEPT to the third author. The research activities of the ICCAS are funded by the *German Federal Ministry for Education and Research (BMBF)* and the *Saxon Ministry of Science and Fine Arts (SMWK)* in the scope of the initiative *Unternehmen Region* with the grant numbers 03 ZIK 031 and 03 ZIK 032. Parts of this work were carried out in the framework of a project supported by the Deutsche Forschungsgemeinschaft (DFG) (Priority Programme 1124, PR 660/3-1).

REFERENCES

1. Klein, S., Strauß, G., Hertel, I., Burgert, O., Tietjen, C., Preim, B. Dietz, A.: „Der Einfluss der Visualisierung von dreidimensionalen radiologischen Befunden in der Halschirurgie". *77. Jahresversammlung der Deutschen Gesellschaft für Hals-Nasen-Ohren-Heilkunde, Kopf- und Hals-Chirurgie*, rheinware Verlag, 2006, p. 151.
2. Velichkovsky, B.M., Pomplun, M. and Rieser. H. (1996). Attention and communication: Eye-movement-based research paradigms. In: W.H. Zangemeister, S. Stiel & C. Freksa (Eds.). *Visual attention and cognition* (pp. 125-154). Amsterdam and New York: Elsevier Science.
3. Mello-Thoms, C., Britton, C., Abrams, G., Hakim, C., Shah, R., Hardesty, L., Maitz, G., and Gur, D. (2003). Head-mounted versus remote eye tracking of radiologists searching for breast cancer: A comparison. *Academic Radiology, 13*(2), 203-209.
4. Nodine, C.F. and Kundel, H.L. (1987).Using eye movements to study visual search and to improve tumor detection. *RadioGraphics 7*(6) 1241-1250.
5. Krupinski, E.A. Nodine C.F. and Kundel HL (1993). A perceptually based method for enhancing pulmonary nodule recognition. *Investigative Radiology, 28*, 289-294.
6. Krupinski, E.A, Nodine, C.F. and Kundel, H.L. (1993). Perceptual enhancement of tumour targets in chest X-ray images. *Perception and Psychophysics, 53*, 519-526.
7. Krupinski, E.A., Berger, W.G., Dallas, W.J. and Rohrig, H. (2003). Searching for nodules: What features attract attention and influence detection? *Academic Radiology, 10*, 861-868.
8. Rötting, M. (2001). *Parametersystematik der Augen- und Blickbewegungen für arbeitswissenschaftliche Untersuchungen*, Dissertation, RWTH Aachen: Shaker Verlag.
9. Velichkovsky, B.M., Baccino, T., Cornelissen, F.W., Geusebroek, J.M., Hansen, L.K., Hari, R., Lőrincz, A., Nicu Sebe, Pannasch, S. and Walter, H., (2006). *PERCEPT: Perceptual Consciousness – Explication and Testing*. EU Commission, Brussels (www.nest-percept.eu).
10. Krüger, A., Tietjen, C., Hintze, J., Preim, B., Hertel, I., Strauß, G., *Interactive Visualization for Neck Dissection Planning*. In IEEE/Eurographics Symposium on Visualization (EuroVis), p. 295-302, 2005.
11. Radermacher, K., Pichler, C.v., Fischer, S. and Rau, G. (1998). *3D-Visualisation in surgery*. Aachen: Helmholtz-Institute / mediTEC.
12. Velichkovsky, B.M., Joos, M., Helmert, J.R., and Pannasch, S. (2005). Two visual systems and their eye movements. *CogSci 2005: Proceedings of the XXVII Conference of the Cognitive Science Society* (pp. 2283-2288). Mahwah, NJ: Lawrence Erlbaum.
13. Abler, B., Walter, H. and Erk, S. (2005). The neural correlates of frustration. *Neuroreport 16*, 669-672
14. Erk, S., Spitzer, M., Wunderlich, A.P., Galley, L. and Walter, H. (2002). Cultural objects modulate reward circuitry. *Neuroreport 13*, 2499-2503.