



The Visualisation of Eye-tracking Scanpaths: What can they tell us about how Clinicians View Electrocardiograms?

Document Version

Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

Citation for published version (APA):

Davies, A., Vigo, M., Harper, S., & Jay, C. (2016). *The Visualisation of Eye-tracking Scanpaths: What can they tell us about how Clinicians View Electrocardiograms?*. Paper presented at ETVIS 2016 - Second Workshop on Eye Tracking and Visualization, Baltimore, Maryland, United States. <http://www.vis.uni-stuttgart.de/etvis/>

Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.



The Visualisation of Eye-tracking Scanpaths: What can they tell us about how Clinicians View Electrocardiograms?

Alan Davies, Markel Vigo, Simon Harper, and Caroline Jay

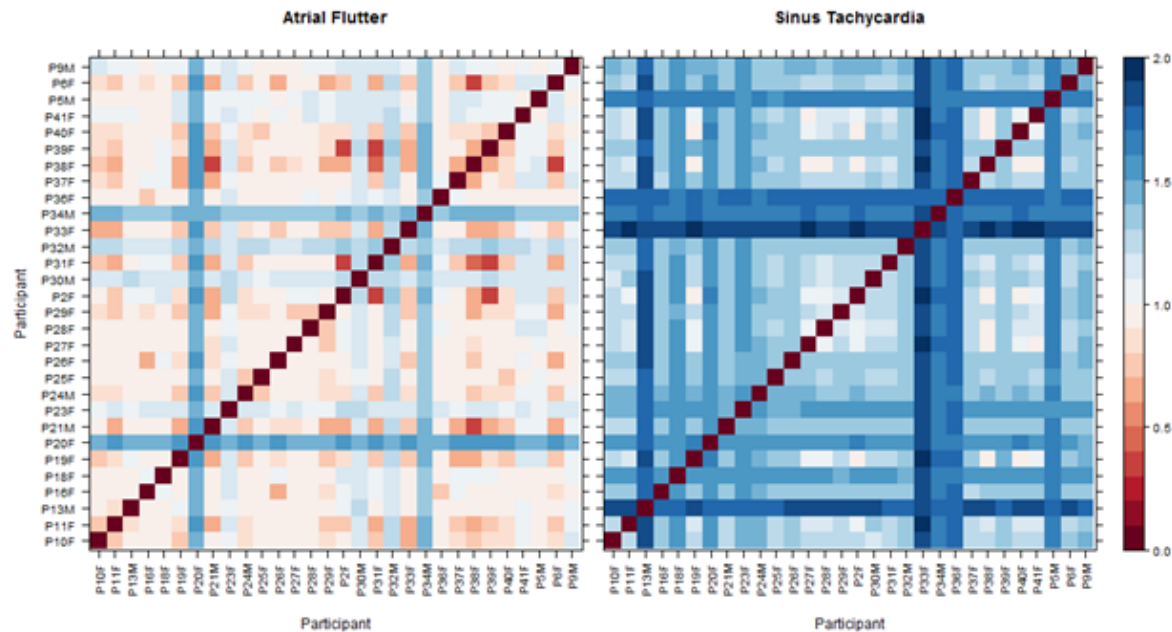


Fig. 1. Stimuli with the most and least similar scanpaths comparing all participants scanpath sequences with each other using the Levenshtein distance (red = more similar, blue = less similar)

Abstract— This paper describes the use of the Levenshtein distance and nearest neighbour index to visualise and analyse differences in eye-tracking scanpaths applied to the field of electrocardiology. Data was obtained from clinicians as they interpreted 12-lead electrocardiograms (ECGs). The main aim of the work is provide methods of visualising the differences between multiple participants scanpaths simultaneously. Allowing us to answer questions such as, do clinicians fixate randomly on the ECG, or do they apply a systematic approach? Results indicate that practitioners have very different search strategies applied to the majority of stimuli. The distribution of fixations is not random and tends towards clustering with all stimuli. The differences between practitioners are likely to be the result of different training, clinical role and expertise.

Index Terms—eye-tracking, visualisation, Levenshtein, scanpaths, electrocardiogram, ECG, EKG

1 INTRODUCTION

Eye tracking has been used to cast light on human-data interaction in a number of medical fields, most notably radiology [17, 20], but this analysis often provides a primarily qualitative interpretation of the diagnostic process. Here we report on some exploratory work used to analyse clinician’s visual behaviour in the medical sub-domain of electrocardiology. This field is particularly interesting, as electrocardiogram (ECG) data consists of signal data from 12 sources, which are presented in sections on a single output (further details of this are provided in 1.1). These sections form pre-existing ‘Areas of Interest’ (AOIs) which it is possible to interrogate from a quantitative perspective.

We examine clinicians’ scanpaths as they interpret ECGs, considering their transition behaviour between the leads. By calculating and visualising the Levenshtein distance and nearest neighbour index we shed light on the extent to which clinicians apply a consistent systematic approach to ECG interpretation.

1.1 Electrocardiology

The 12-lead electrocardiogram (ECG) is a graphical representation of the electrical activity generated by the heart [8]. Due to its relatively low cost and increasing availability it is one of the most frequently performed diagnostic medical tests carried out in a wide variety of clinical settings [8]. The ECG is printed onto gridded paper with the x-axis representing time in seconds and the y-axis representing amplitude in millivolts [6]. The ECG waveform (figure 2) consists of a series of waves, intervals and segments representing depolarisation and repolarisation of various parts of the cardiac conduction system [8, 25]. The ECG is also split into 12 different regions, termed leads. Leads I, II, III, aVR, aVL, aVF “view” the coronal/frontal plane, whereas leads V1-V6 view the transverse plane of the heart. The waveforms appear differently in the different leads as the cardiac impulse is deflected

• Alan Davies, Markel Vigo, Simon Harper and Caroline Jay are from the University of Manchester, Manchester UK. E-mails: alan.davies-2@manchester.ac.uk, or firstname.lastname@manchester.ac.uk.

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxx/TVCG.201x.xxxxxx

relative to its direction of travel towards or away from the positive and negative poles of the surface electrodes (cables attached to the body surface to detect the heart’s electrical activity) [9].



Fig. 2. The ECG waveform (lead II)

ECG interpretation is a complex task undertaken by many healthcare professionals including doctors of all specialities, nurses, physiologists/technicians and paramedics. Incorrect interpretation can lead to no or incorrect treatment being given, potentially leading to injury or death [15]. Current automated or computerised interpretation (CI) software is frequently inaccurate [2], and human interpreters are still considered the “gold standard” of interpretation [24]. This work aims to explore human visual behaviour in relation to ECG interpretation using eye-tracking. Understanding how human visual behaviour can impact on ECG interpretation could represent a pivotal stage in improving clinically reliable ECG interpretation software. This work consists of a fusion of different disciplines, including medicine, psychology and computing. The first stage of this process – reported on here – involves visualising eye movement data to aid hypothesis generation.

1.2 Related work

Scanpath analysis is widely used in eye-tracking research [14]. Scanpaths can be considered in terms of the sequence of AOIs (areas of interest defined by the researcher) that a participant visits, analysed with techniques such as the Levenshtein distance or in terms of the spatial positions/alignment of fixations (vector sequence alignment). Methods can also include temporal aspects, such as fixation duration and saccadic amplitude (vector strings) [14]. Scanpath analysis provides insight into the cognitive processes of users interacting with a visual stimulus, as eye-tracking has been linked to decision making [11].

Scanpath comparison measures provide an estimation of similarity between ≥ 2 scanpaths [14]. The Levenshtein distance is a metric used for calculating the differences between two string sequences. It applies a cost (numerical penalty) for the number of insertion, deletion and substitution operations needed in order to ensure the strings contains the same tokens in the same order [19]. Levenshtein distance remains one of the most commonly used methods for comparing scanpaths [14, 18] and has been used across a multitude of domains, including reasoning about the mental state of others [21] and website scanning [23]. Alternative string edit distances include Longest common subsequence (LCS), and Damerau-Levenshtein and Hamming distance [18]. The Levenshtein distance has also been enhanced and adapted: [12] applied the Levenshtein distance augmented by the Needleman-Wunsch approach, which allows the user to define cost functions, to improve the diagnostic/evaluation methods of attention disorders classification in the medical domain. Methods of scan path visualisation include the Voronoi method, which uses Voronoi diagrams to divide areas into cells with the distribution of cells correlating to the fixation density distributions. The method is essentially a spatial method that is analogous to fixation clustering [22]. Dotplots also provide a visual method for viewing scanpath similarities and can be used for exploration and validation purposes [13]. Here, we trial a method for visualising similarities and differences between individuals’ scanpaths while they interpret ECGs.

2 METHOD

Thirty one participants whose clinical role includes regularly interpreting ECGs (Table 1) viewed eleven ECGs taken from open access on-line libraries ¹ and displayed in random order on a computer screen.

¹<http://lifeinthefastlane.com/ecg-library/> and www.emedu.org/ecg_lib/index.htm

Participant clinical role	n	%
Cardiac physiologist / technician	20	62.5
Student cardiac physiologist / technician	4	12.5
Registrar	1	3.1
Cardiac registrar	2	6.3
Medical student	1	3.1
Senior nurse	1	3.1
Consultant	1	3.1
FY2 doctor	1	3.1
Health care assistant	1	3.1
Gender		
Male	8	26
Female	23	74

Table 1. Participant demographics

Each participant had as long as they required to interpret each ECG. Once they had reached an interpretation they were asked to speak it aloud. This was recorded via a Zoom H2 Handy Recorder. Participants moved through the ECGs by clicking a mouse button to proceed. Eye-tracking data was captured for each participant with Tobii 1750 and X2-60 eye trackers. Tobii software allows for the creation of areas of interest (AOIs); these are areas that encapsulate some entity of the stimulus that researchers are interested in. AOIs were mapped onto each of the leads and rhythm strip (an existing lead repeated for a longer duration usually located at the bottom of the ECG). The AOIs were labelled A to M. A full description of the study protocol, stimuli and resulting data are available from our data repository ².

3 ANALYSIS

Determining the level of similarity between eye-movements across different participants is important [4]. Many of the standard techniques such as, heat/focus maps and gaze plots do not display the sequential/temporal nature of these movements [4], or allow multiple subject comparisons, without introducing excess visual complexity with higher numbers of participants. Scanpaths represent the sequence of visual transitions around a stimulus. Gazeplots represent this sequence of fixations made by participants viewing a stimulus, enabling a visual comparison between participants to be made. This can, however, become overly complex or meaningless when attempting to compare larger groups, or all of the participants in a study (figure 3).

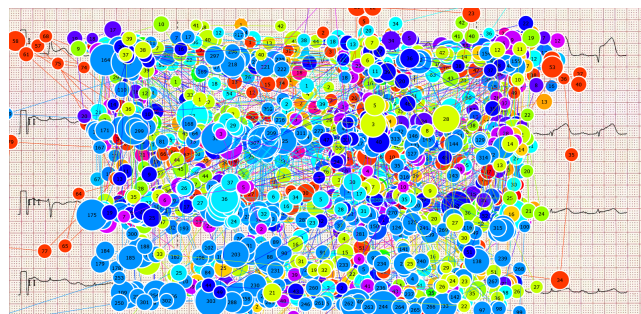


Fig. 3. Scan paths overlaid on stimulus (all participants)

Scan paths can also be represented as a set of characters, known as ‘strings’, detailing the sequence in which a participant visited each AOI, as demonstrated in this example from two of the study participants:

$$P_2 = \{M, F, E, M, I\}$$

$$P_{23} = \{F, D, G, C, M, I, H, J, G, H\}$$

Differences in fixation duration, or total time spent viewing the stimulus could identify similarities between participants, however this is limited to presenting temporal data and does not capture the

²<http://iam-data.cs.manchester.ac.uk/investigations/12>

similarity in attention that certain areas of the stimulus present, or the visual transitions around the stimulus that may be required when cross referencing areas of clinical importance. To this end the Levenshtein distance was used to compute the similarity of the participants' scanpath strings with one another. The distance is defined as the minimal number of operations needed to transform one string into another using insertion, deletion and substitution operations [14, 19]. Prior to applying the Levenshtein distance the scanpath strings were truncated removing consecutive tokens to focus on AOI visited and removing fixation frequency, i.e. a scanpath string consisting of $\{M, M, M, B, B, A, B, C\}$ would become $\{M, B, A, B, C\}$. This was done to examine the sequence of transitions between leads, considering dwells rather than fixations. Each of the participants' scanpaths were compared against all the other participants' for each stimulus (ECG) and the results were displayed using a matrix to allow for rapid visual comparison. A log transformation was applied to facilitate comparison across the stimuli. Figure 1 shows the stimuli that represent the most and least similar scanpaths. The higher the value (indicated by increasing 'blueness') the more different the scanpaths. A darker shade of red conversely shows a greater similarity in the scanpaths. This method of visualisation also makes it easier to spot outliers.

$$NNI = \frac{d(NN)}{d(ran)} \quad (1)$$

$$where : d(NN) = \sum_{i=1}^N \left[\min \left(\frac{d_{ij}}{N} \right) \right] \quad (2)$$

$$and : d(ran) = 0.5 \sqrt{\frac{A}{N}} \quad (3)$$

$$and : d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (4)$$

Equation 1: Nearest neighbour index (NNI)

We used the Nearest Neighbour Index (NNI), which provides a ratio of departure from randomness for a distribution pattern [5], to determine if the fixations are regularly spaced or clustered. The technique does not offer any explanation of the patterns, but a random result would indicate that there is no underlying pattern at all. In this study we may infer that the distribution patterns indicate either a systematic approach to searching the leads of the ECG (uniform/regular), or perhaps a focus on certain specific areas (clustered), which we hypothesise are likely to represent the salient features of the condition. The NNI is a quantifiable measure of a distribution pattern and is a frequently used distance statistic that allows for the analysis of geospatial distributions, defining the dispersion on a continuum as being clustered, random or regular. Values include a lower bound of 0 (clustered) and an upper bound of 2.15 (regular) where values of 1 are considered random (equation 1) [10]. The Euclidean distance (equation 1.4) was used to compute the straight line distance between fixations.

4 RESULTS

4.1 Levenshtein distance

The average log transformed Levenshtein distance shows that the most similar scanpath sequences are found in the atrial flutter (average Levenshtein distance = 0.99) and torsades de points (0.94) ECGs (figures 5 and 4). The least similar scanpaths are found in the sinus tachycardia ECG (1.42). Figure 5 displays the mean log transformed Levenshtein distance for each of the ECGs, providing a comparison of the similarity of the scanpaths across the stimuli. The implication of this is that the participants visual search strategies are very different from one another in the majority of the stimuli, with the exception of two aforementioned ECGs.

4.2 Nearest Neighbour Index (NNI)

Results show that the fixations are not randomly distributed: all 11 stimuli show a tendency towards clustering with values between 0 and 1, and an average NNI of 0.375 (see Figure 6).

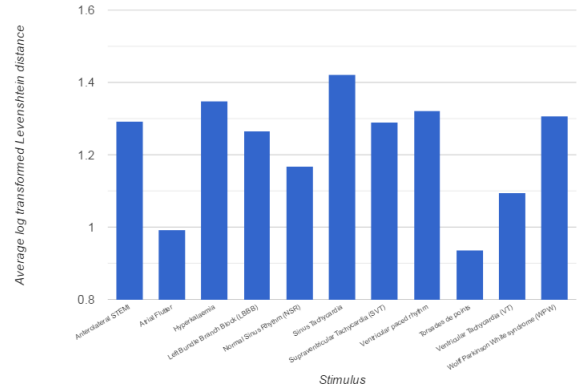


Fig. 5. Average log transformed LD (all stimuli)

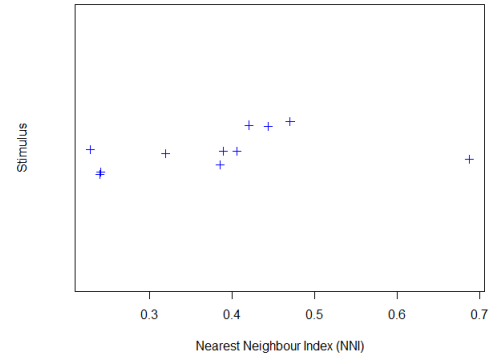


Fig. 6. Nearest Neighbour Index (all stimuli)

The Hyperkalaemia stimulus seen on the far right of figure 6, with an NNI of 0.6877, was the least relatively clustered stimulus. This is also the condition that had the largest scanpath sequences.

5 DISCUSSION

The results show large differences between the participants' scanpaths, which is indicative of differing search strategies. This difference could be attributable to the disparate backgrounds of the participants. There are many different methods of teaching ECG interpretation that vary in approach and duration [1]. These methods also differ between countries and institutions as well as the medical discipline that the practitioner belongs to [16]. Using a matrix to visualise the similarities/differences between the scanpaths with the Levenshtein distance is a helpful initial way of gaining a comparative overview of multiple participants in a study, and locating outliers who have markedly different or similar scanpaths. Comparing the plots of different stimuli may also indicate some similarity/difference that originates from some aspect of the stimulus itself. Scanpath analysis suffers from some limitations, including the issue of scanpath length, with considerably different lengths confounding alignment calculations [13]. It should also be noted that visual behaviour is very rich, and 'naive' scanpath analysis will not tell the whole story. Future work will focus on refining this approach, by considering visual transitions between leads, which is discussed in more detail in other work [7], and will also consider how factors such as accuracy of interpretation affect the results.

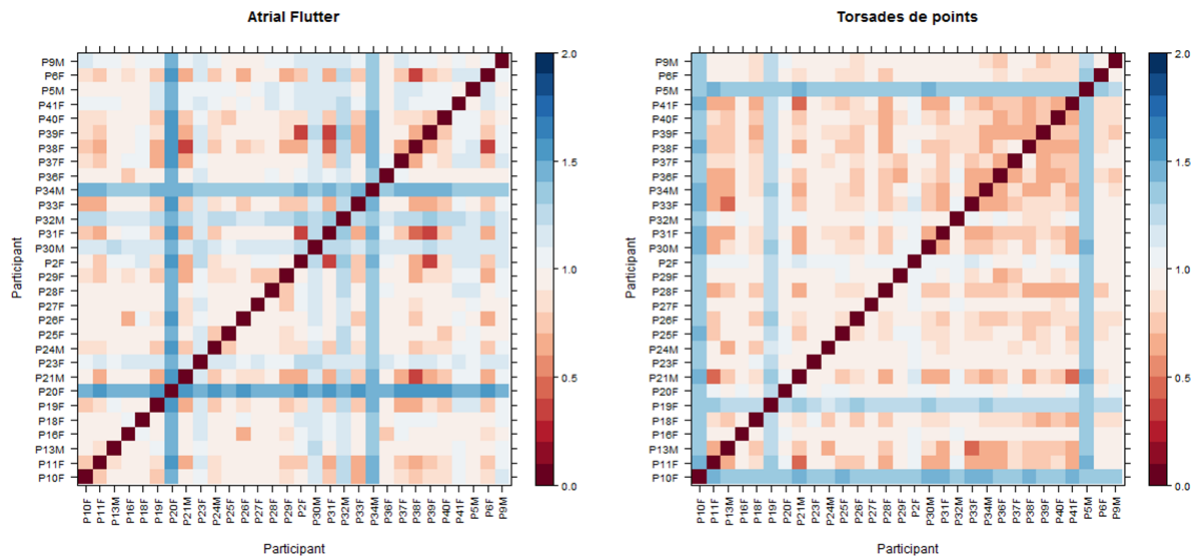


Fig. 4. Stimuli with the most similar scanpath distances (lowest mean Levenshtein distance)

The NNI has been used as a measure of mental workload, where a more random pattern was indicative of a higher workload when applied to dynamic scenes [10]. The results of our study show a tendency toward clustering with 10/11 of the stimuli having an NNI of <0.5 . This is in keeping with previous work that demonstrates people tend to focus on some leads more than others [3]. This indicates that participants were drawn toward specific features, possibly the lead or a component of the waveform that displays features of the ECG abnormality. This may also be the case regardless of making a correct or incorrect interpretation as a participant may notice an abnormal feature without necessarily understanding its significance. Eye tracking data is frequently used to augment usability studies [11]. The small sample sizes frequently used in usability studies coupled with the richness of eye-tracking data can make analysis of datasets such as the one used in this study challenging and often non amenable to traditional statistical approaches. The techniques described in this paper go some way toward providing a quantitative approach for exploration of this type of data, and we therefore anticipate they will have a scope wider than the ECG sub-domain, as they provide a means of understanding whether individuals are employing a systematic approach, or have some intrinsic similarity in their visual behaviour.

6 CONCLUSIONS AND FUTURE WORK

The methods presented here offer a way of exploring and visualising the visual behaviour of practitioners viewing ECGs. They allows us to visualise differences in scanpaths that can indicate different search strategies, which may result from different training or experience. With a larger sample size (>30 per group) the NNI could be calculated for the groups of practitioners making correct and incorrect interpretations to identify differences in fixation distribution across the stimulus. A weighted distance metric could also be introduced to incorporate the effect of time spent viewing the areas, as well as transitions between them. The techniques provide a way of viewing the similarities and differences in multiple scanpaths and stimuli simultaneously, providing a quantifiable measure of difference without increasing visual complexity.

ACKNOWLEDGMENTS

EPSRC: EP/K502947/1 and EP/L504877/1

REFERENCES

- [1] G. Alinier, R. Gordon, C. Harwood, and W. B. Hunt. 12-Lead ECG training: The way forward. *Nurse Education Today*, 26:87–92, 2006. doi: 10.1016/j.nedt.2005.08.004
- [2] D. Anh, S. Krishnan, and F. Bogun. Accuracy of electrocardiogram interpretation by cardiologists in the setting of incorrect computer analysis. *Journal of electrocardiology*, 39(3):343–5, 7 2006. doi: 10.1016/j.jelectrocard.2006.02.002
- [3] R. R. Bond, T. Zhu, D. D. Finlay, B. Drew, P. D. Kligfield, D. Guldenring, C. Breen, a. G. Gallagher, M. J. Daly, and G. D. Clifford. Assessing computerized eye tracking technology for gaining insight into expert interpretation of the 12-lead electrocardiogram: an objective quantitative approach. *Journal of electrocardiology*, 47(6):895–906, 2014. doi: 10.1016/j.jelectrocard.2014.07.011
- [4] J. Cirimele, J. Heer, and S. K. Card. The VERP Explorer : A Tool for Exploring Eye Movements of Visual-Cognitive Tasks Using Recurrence Plots. 2014.
- [5] P. J. Clark and F. C. Evans. Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*, 35:445–453, 1954.
- [6] G. D. Clifford, F. Azuaje, and P. E. McSharry. *Advanced Methods and Tools for ECG Data Analysis*. Artech house, Inc, Boston, 2006.
- [7] A. Davies, C. Jay, S. Harper, and A. Brown. Computational Methods for Analysis of Visual Behavior using Eye-tracking. In A. Spink, ed., *Proceedings of Measuring Behavior 2016, (Dublin, Ireland, 25-27 May 2016)*, number May 2016, pp. 25–27. Dublin, Ireland, 2016.
- [8] A. Davies and A. Scott. *Starting to read ECGs: The Basics*. Springer-Verlag, London, 2014. doi: 10.1007/978-1-4471-4962-0
- [9] A. Davies and A. Scott. *Starting to read ECGs: A Comprehensive Guide to Theory and Practice*. Springer-Verlag, London, 2015. doi: 10.1007/978-1-4471-4965-1
- [10] F. Di Nocera, M. Terenzi, and M. Camilli. Another Look at Scanpath: Distance to Nearest Neighbour As a Measure of Mental Workload. *Developments in Human Factors in Transportation, Design & Evaluation*, (July 2015):295–303, 2006.
- [11] C. Ehmke and S. Wilson. Identifying web usability problems from eye-tracking data. In *HCI 2007*, pp. 119–128, 2007.
- [12] F. Galgani, Y. Sun, P. L. Lanzi, and J. Leigh. Automatic Analysis of Eye Tracking Data for Medical Diagnosis. 2009.
- [13] J. H. Goldberg and J. I. Helfman. Scanpath clustering and aggregation. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications - ETRA '10*, p. 227, 2010. doi: 10.1145/1743666.1743721
- [14] K. Holmqvist, M. Nystrom, R. Anderson, R. Dewhurst, H. Jarodzka, and J. Van de Weijer. *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press, New York, 2011.
- [15] H. Holst, M. Ohlsson, C. Peterson, and L. Edenbrandt. A confident

- decision support system for interpreting electrocardiograms. *Clinical Physiology*, 19(5):410–418, 1999.
- [16] a. H. Kadish, a. E. Buxton, H. L. Kennedy, B. P. Knight, J. W. Mason, C. D. Schuger, C. M. Tracy, W. L. Winters, a. W. Boone, M. Elnicki, J. W. Hirschfeld, B. H. Lorell, G. P. Rodgers, and H. H. Weitz. ACC/AHA clinical competence statement on electrocardiography and ambulatory electrocardiography: A report of the ACC/AHA/ACP-ASIM task force on clinical competence (ACC/AHA Committee to develop a clinical competence statement on electrocardiography and am. *Circulation*, 104(25):3169–78, 12 2001.
- [17] E. A. Krupinski, N. F. Calvin, and H. K. L. Enhancing recognition of lesions in radiographic images using perceptual feedback. *Optical Engineering*, 37(3):813–818, 2013.
- [18] O. Le Meur and T. Baccino. Methods for comparing scanpaths and saliency maps: strengths and weaknesses. *Behavior research methods*, 45(1):251–266, 2013. doi: 10.3758/s13428-012-0226-9
- [19] V. I. Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals, 1966. doi: citeulike-article-id:311174
- [20] D. Litchfield, L. J. Ball, T. Donovan, D. J. Manning, and T. Crawford. Learning from others: effects of viewing another person’s eye movements while searching for chest nodules. *Medical Imaging*, 6917:691715–691715, 3 2008. doi: 10.1117/12.768812
- [21] B. Meijering, H. van Rijn, N. A. Taatgen, and R. Verbrugge. What Eye Movements Can Tell about Theory of Mind in a Strategic Game. *PLoS ONE*, 7(9), 2012. doi: 10.1371/journal.pone.0045961
- [22] E. a. B. Over, I. T. C. Hooge, and C. J. Erkelens. A quantitative measure for the uniformity of fixation density: The Voronoi method. *Behavior research methods*, 38(2):251–261, 2006. doi: 10.3758/BF03192777
- [23] B. Pan, H. a. Hembrooke, G. K. Gay, L. a. Granka, M. K. Feusner, and J. K. Newman. The determinants of web page viewing behavior: an eye-tracking study. *Proceedings of the ETRA '04 Symposium on Eye Tracking Research and Applications*, 1(212):147–154, 2004. doi: 10.1145/968363.968391
- [24] S. M. Salerno, P. C. Alguire, and H. S. Waxman. Competency in Interpretation of 12-Lead Electrocardiograms: A Summary and Appraisal of Published Evidence. *Annals of Internal Medicine*, 138:751–760, 2003. doi: 10.1016/S1062-1458(03)00283-6
- [25] G. Wagner. *Marriott’s Practical Electrocardiography*. Lippincott Williams & Wilkins, 11th ed., 2008.