Monitoring Dementia with Automatic Eye Movements Analysis

Yanxia Zhang, Thomas Wilcockson, Kwang In Kim,
Trevor Crawford, Hans Gellersen, and Pete Sawyer

Lancaster University, U.K.
{yazhang,t.wilcockson,k.kim,t.crawford,h.gellersen,p.sawyer}@lancaster.ac.uk

Abstract. Eye movement patterns are found to reveal human cognitive and mental states that can not be easily measured by other biological signals. With the rapid development of eye tracking technologies, there are growing interests in analysing gaze data to infer information about people' cognitive states, tasks and activities performed in naturalistic environments. In this paper, we investigate the link between eye movements and cognitive function. We conducted experiments to record subject's eye movements during video watching. By using computational methods, we identified eye movement features that are correlated to people's cognitive health measures obtained through the standard cognitive tests. Our results show that it is possible to infer people's cognitive function by analysing natural gaze behaviour. This work contributes an initial understanding of monitoring cognitive deterioration and dementia with automatic eye movement analysis.

Keywords: Machine learning, eye movements analysis, health monitoring, dementia, cognitive function

1 Introduction

Healthy cognitive function is essential to live an independent life. Decline in cognitive health can impact how we perform in daily activities, including walking, making food and interacting with people. A major cause of cognitive decline is dementia, a condition that currently affects around one in six people at the age of 80. Increasing life expectancy means that the number of people who develop dementia will increase. Taking the UK as an example, the number of people living with the condition is predicted to increase from the current figure of 850,000 to over 2m by 2051 [17].

Although dementia is currently irreversible and ultimately fatal, obtaining an early diagnosis can help maintain quality of life by treating debilitating side effects, such as depression. Moreover, when improved therapies do eventually become available, it is likely that they will have to be administered before the damage to the brain becomes so severe as to render the therapy ineffective. Currently, diagnosis of dementia or of its harbinger, Mild Cognitive Impairment (MCI), is usually performed using paper-based cognitive tests such as the

Montreal Cognitive Assessment (MoCA [18]). These are designed to be administered in a clinical setting such as a memory clinic but this can be stressful for the subject and yield poor ecological validity. Worse, many subjects don't refer themselves for a health check until the disease is well advanced. There is therefore a strong interest in developing new techniques for detecting cognitive decline that don't suffer from these disadvantages.

One strand of work seeks to test for deficits in the same cognitive domains (memory, executive function, motor control and so on) that are tested by the paper tests, for example, using everyday computer tasks as proxies for tasks in the tests [15]. However, our work builds upon studies that have shown that eye movements are a bio-marker for dementia [8, 7, 1].

We are engaged on a programme of research in which our goal it to develop ambient eye-tracking systems for the detection of cognitive decline. This work seeks to identify a set of eye movement features that are correlated to variations in cognitive capability among people with cognitive impairment and healthy people. This will deepen our understanding of the link between cognitive health and eye movements and provide insights that we can exploit in the design of our envisioned ambient eye-tracking system.

The primary contribution of this paper is to introduce a computational method that analyses natural gaze behaviour automatically to predict cognitive function. Following a brief review of the literature, we describe an experiment with 15 participants to record their gaze data during video watching. We then report the computed statistical eye movement features and correlate them with the cognitive assessment scores obtained through paper cognitive tests. The paper concludes with a summary of the findings and discuss the extent to which our results could be used for health monitoring.

2 Related Work

Eye movements have been shown to reflect a combination of top-down cognitive processes (e.g., the observer's task, interest and goals) and bottom-up perceptual processes (e.g., influences of low level image properties) [13]. A growing body of evidence supports the use of eye movements for predicting physical activities being performed [5] and the human performers' internal mental states [4]. Eye movements are also found to provide a sensitive marker of cognitive change or deterioration [8, 7]. Patients with dementia lose the efficient control of attention and have an impairment of inhibitory control and error-correction that exceeds the effects of normal ageing [8]. In anti-saccade tasks, error frequency is correlated with dementia severity, demonstrating a potential for eye tests that provides quantitative measures for dementia diagnosis [7].

2.1 Decoding Mental and Cognitive States

A number of studies have examined specific types of eye movements and their relation to variations in mental and cognitive states. *Di Stasi et* al. investigated saccade as a diagnostic measure of mental workload [9]. They tested 18 subjects in a virtual driving task with three complexity levels and found that saccadic peak velocity decreased as the mental workload increased. Their results suggest that saccadic peak velocity could be a useful diagnostic index for the assessment of operators' mental workload and attentional state in hazardous environments. *Schleicher et* al. (2008) [19] examined changes in a variety of oculomotoric variables (e.g., blink duration) as a function of increasing sleepiness.

Recently, there has been an increasing amount of work looking into machine learning methods to automatically decode mental and cognitive states [4]. *Steichen et* al. analysed eye gaze patterns in interactive visualisation tasks using a number of classification methods [20].Their results showed that using simple machine learning on eye tracking metrics can infer a number of task and user characteristics. A number of studies demonstrated the possibility of using eye movement features to classify mental states in scene viewing tasks [14, 12, 16]. *Kardan et* al. recorded eye movements from 72 participants while performing three tasks: visual search, scene memorization, and aesthetic preference [16]. They used statistical features (mean, standard deviation, and skewness) of fixation durations and saccade amplitudes, as well as the total number of fixations. Their results showed that eye movement distributional properties can classify mental states both within and across individuals. Eye movements are also found to be signatures of implicit navigational and information search intention [14], interaction intents in command issuing [2].

2.2 Health Applications

Eye movement studies have shown that people with neurological conditions exhibit abnormal viewing patterns [1]. With the advancement of computational tools and eye tracking hardware, there is growing interest in using eye movements for health applications.

Benson et al. investigated which eye movement tests (smooth pursuit, fixation stability, and free-viewing tasks) alone and combined can best discriminate Schizophrenia cases from control subjects [3]. Their results showed that a boosted tree model achieved perfect separation of the 88 training cases from 88 control subjects. Its predictive validity on retest assessments and novel cases and control subjects was 87.8%. However, a probabilistic neural network model was superior and could discriminate all cases from controls with near perfect accuracy at 98.3% when evaluated on the whole data set of 298 assessments.

Tseng et al. devised a high-throughput, low-cost method to classify clinical populations where participants simply watched television [21]. Based on a computational model of visual attention, they extracted 224 quantitative features from the patients and controls' eye tracking data. Using machine learning in a work flow inspired by microarray analysis, they identified critical features that differentiate patients from control subjectswhich classified Parkinson's disease versus age-matched controls with 89.6 % accuracy (chance 63.2 %), and attention deficit hyperactivity disorder (ADHD) versus fetal alcohol spectrum disorders (FASD) versus control children with 77.3 % accuracy (chance 40.4

%). Similarly, *Crabb et* al. tested the hypothesis that age-related neurodegenerative eye disease (e.g., glaucoma) can be detected by examining patterns of eye movement recorded while a person naturally watches a movie [6]. They proposed a novel method to generate saccade density maps from scanpaths of eye movements recording and used kernel principal component analysis (KPCA) for feature extraction. They found that the generated saccadic maps can contain a signature of vision loss which can separate patients from healthy peers at reasonable accuracy. These results demonstrated that by automatic analysing eye movement patterns during visual activities (e.g., scene viewing, video watching) can access individual's cognitive health.

3 Methods

Our goal is to correlate eye movement patterns during video watching with an individual's assessment score from a battery of cognitive tests. We are building on what is known about the eye movement features of cognitively impaired people to see if they can be artificially stimulated and observed by video-watching tasks. In this experiment, we aim to find out which features are important in predicting the cognitive scores.

3.1 Participants

This study collected data from 15 participants (8 female and 7 male) with 9 older control participants (mean 66.11, std 9.57), 3 young controls (mean 26.33, std 4.04) and 3 MCI patients (mean 71.67, std 2.52). MCI is interesting because the sufferer will exhibit cognitive decline greater than that expected for their age, but which is not yet so severe as to significantly inhibit their day-to-day functioning. Someone with MCI is at high risk of developing dementia within 5 years. All participants participated voluntarily. Older adult controls were over 55 years old and recruited from a local church, younger adult controls were recruited from a local university, and MCI participants were recruited from National Health Service (NHS) Memory Services.

Fig. 1: The experiment recorded participants' eye movements using a remote eye tracker while watching short video clips.

3.2 Apparatus

An EyeLink Desktop 1000 eye-tracker (SR Research Ltd., Ontario, Canada) was used at 500 Hz. Participants sat approximately 55 cm away from the monitor (60 Hz) (see Figure 1). Their dominant eye was determined using the Miles test and tracked accordingly. Experimenter Builder software Version 1.10.1630 was used to control the stimulus events during the eye-tracking task.

3.3 Stimuli and Tasks

Each participant in the experiment underwent a memory test (FCSRT-IR - see the end of this sub-section) and then performed an eye movement task that required them to watch four short videos. Each video lasted 40 seconds. Three of these videos were viewed on three occasions each. The fourth video was only viewed once and required participants to fixate on one object for the duration of the video. Prior to the viewing of each video the participant was given instructions related to the video. In the *free view* session a participant was asked to freely process the video in order to obtain a measure of attention-catching objects that are highly salient. In the second two *instructed* sessions a specific question was asked which was designed to direct the top-down control of eye gaze to non-salient objects that are not necessarily attention-catching. The entire experiment lasted about an hour. The participants watched the videos in a pre-defined order, as we are only interested in eye movement patterns irrespective of the viewing content. In total, we collected 10 video trials per participant. Table 1 describes the instructions on the videos viewed on three occasions.

Table 1: Videos viewed during the study

Video	Sessions
1: Coronation of the Queen Elizabeth II	i Free view ii How many bald men are in the room? iii What are the colours of the clothes within the room?
2: Neil Armstrong landing on the moon	i Free view ii How many legs has the moon lander? iii Can you see the astronauts' faces?
3: Gordon Brown and family leaving Downing Street after losing the general election in 2010	i Free view ii What are each member of Gordon Brown's family wearing? iii How many windows are there on the buildings?
4: Hovis: an old advertisement for Hovis bread	i Fixate on the boy with the bicycle in each scene.

The participants' answers were recorded. However, we were primarily interested in differences in eye movement patterns toward salient stimuli in the different viewing conditions rather than whether the participant answered the question correctly (in fact, all participants did answer the questions accurately). Across participant groups the eye movement data was investigated in association with performance in the Free and Cued Selective Reminding Test with Immediate Recall (FCSRT-IR [10]). The FCSRT-IR is a measure of memory which is not confounded by normal age-related changes in cognition and has been associated with preclinical and early dementia. Participants are asked to memorise line drawings of easily recognised objects (e.g., grapes) which belong to unique category cues (e.g.,fruit). A measure of free recall and a measure of cued recall is obtained by calculating the correct responses (both out of a total of 48).

3.4 Data Analysis

Feature extraction: Eye movements recorded before and after video watching and during blinks were removed. Because we are interested in eye movement patterns rather than prolonged fixations, we filtered out fixations that are longer than 1500ms and shorter than 10ms and saccades with amplitude larger than 100 degrees. We extracted the distribution features from the fixation and saccade data per trail. These include the mean, standard deviation, skewness (a measure of symmetry) of the fixation duration, saccade amplitude, average and peak velocity, as well as the number of fixations over each video viewing. 13 features were computed from each video trial. We extract these features because [16] showed that distribution properties of fixations are effective features for classifying mental states.

Prediction model: We used the Pearson correlation coefficients to identify important features that are correlated to the memory function as assessed by the FCSRT-IR measure. We further applied three linear regression models: least-squares regression, ridge regression, and LASSO regression [11] to test our hypothesis that automatic analysing eye movement patterns can predict individual's memory capability. If we denote the input feature vector and the corresponding output memory measure as $\mathbf{x} = [x_1, x_2, ..., x_{13}]^\top$ and y, respectively, all three regression models can be expressed as:

$$\hat{y} := f(\mathbf{x}) = \mathbf{x}^\top \mathbf{w}.$$

Least-squares regression f_S is obtained by minimising the training error: For given n training examples of input feature vectors and the corresponding outputs $\{(\mathbf{x}^1, y^1), \ldots, (\mathbf{x}^n, y^n)\}$, f_S minimises

$$\mathcal{E}[f] = \sum_{i=1}^{n} (y^i - f(\mathbf{x}^i))^2.$$

In general, simple least-squares regression can overfit to data, i.e., it fits perfectly to the training data but generalizes poorly. Ridge regression overcomes this problem by introducing a *Tikhonov regularization*: The solution f_R minimizes the sum of squared error and a ridge penalty which measures the smoothness of a function: $\mathcal{E}[f] + \lambda_R \|f\|_2^2$ where λ_R is a hyper-parameter. Finally LASSO replaces the L2 panalty in the ridge regression with $L1$ regularizer: LASSO estimator f_L minimises [11]

$$\mathcal{L}[f] = \sum_{i=1}^{n} (y^i - f(\mathbf{x}^i))^2 + \lambda_L \|f\|_1, \qquad (1)$$

By trading empirical risk off with L1 penalty, LASSO tends to set some elements of the resulting coefficient vector $\mathbf{w}_L$ zero. This leads to automatic feature selection as the features corresponding to non-zero coefficients in $\mathbf{w}_L$ can be regarded as more *influential* in constructing the prediction.

To evaluate the regression models, training was applied on 70% of the trials. We use the remaining 30% of the trials for testing. The model parameters λ in LASSO and Ridge regression were selected with 10-fold cross validations. We conducted 100 iterations of randomly sampling the training and testing sets.

4 Results

We first summarise the FCSRT-IR memory scales (out of 48) across subjects. The statistics show that free recall memory scores in the MCI group (mean = 19.67, std = 10.07) are lower than those in the healthy young (mean = 38.00, std = 4.36) and old control groups (mean=35.33, std=4.27).

4.1 Correlation with Memory Scale

The first goal of our analysis is to estimate the significance of the memory function differences per statistical feature extracted from the eye movements. Table 2 shows a set of correlations of eye movement features to free memory recall scale. Pearson's r ($-1 \leq r \leq 1$) indicates the strength and direction of the correlation, where 1 is total positive correlation, 0 is no correlation and -1 indicates a perfect negative correlation.

There is a significant positive correlation between the skewness of fixation duration and free memory recall scale (r=0.3696, p=0.0125) in the free view condition. The distribution of fixation durations of subjects with high memory scores exhibits a long tail to the right (longer durations). The majority of fixation durations are concentrated to the left (shorter durations). While in the instructed conditions, we find a significant negative correlation between the mean of fixation duration and free memory recall scale (r=-0.2485, p=0.0182). This indicates that the average of fixation time is shorter for subjects with high memory scores.

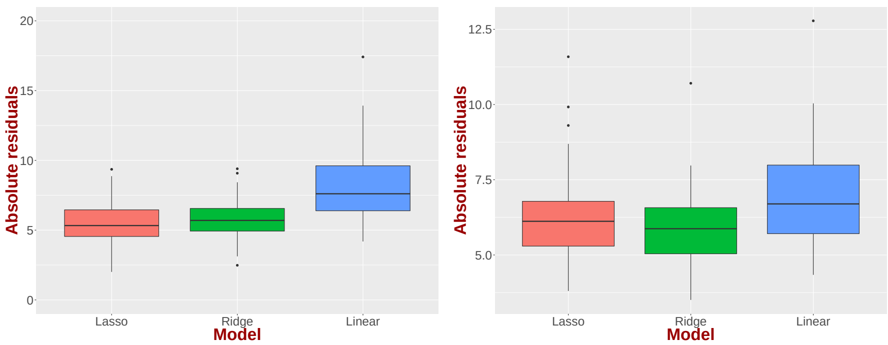
There is a strong significant positive correlation between the mean of saccade average values of velocity in both free view (r=0.4943, p=0.0006) and instructed

Table 2: Correlations between eye movement features and free recall measures

Feature	Free view		Instructed	
	r	p-value	r	p-value
Mean fixation duration	-0.1248	0.4140	−0.2485	0.0182*
Std fixation duration	-0.1228	0.4216	−0.0949	0.3735
Skewness fixation duration	0.3696	0.0125*	−0.0248	0.8166
Fixation count	-0.0685	0.6549	0.0772	0.4697
Mean saccade amplitude	0.1728	0.2564	0.0991	0.3527
Std saccade amplitude	-0.2653	0.0782	−0.0056	0.9571
Skewness saccade amplitude	-0.1439	0.3457	−0.0099	0.9263
Mean saccade average velocity	0.4943	0.0006***	0.3973	0.0001***
Std saccade average velocity	0.0157	0.9182	0.2737	0.0090**
Skewness saccade average velocity	-0.0527	0.7305	0.1240	0.2441
Mean saccade peak velocity	0.0679	0.6578	0.2254	0.0326*
Std saccade peak velocity	-0.0409	0.7895	0.2281	0.0306*
Skewness saccade peak velocity	-0.0073	0.9623	0.1459	0.1700

Notes: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

conditions (r=0.3973, p=0.0001). This shows that subjects with low memory scores showed slower saccade motion on average. This result is also reflected in the instructed condition where the mean saccade peak velocity is found to be positively correlated with the memory scale (r=0.2254, p=0.0326). The standard deviations of the average (r=0.2737, p=0.0090) and peak (r=0.2281, p=0.0306) saccade velocity are positively correlated with the memory recall scale, which indicates that the distribution of the saccade speed is more spread in subjects with low memory score.



(a) Free view (b) Instructed

Fig. 2: Boxplots of the mean absolute residuals across three regression models.

4.2 Automatic Prediction of Memory Capability

The second goal of our analysis is to construct an optimal model based on the input eye movement to predict an individual's memory capability (The memory score has a total of 48). We fit three linear regression models to predict memory score given a set of input features extracted from eye movement recordings. Figure 2 illustrates the boxplots of the average absolute residuals (difference between the predict score and ground truth) for the three regression models in both free view and instructed conditions. Among all three models, lasso achieves the best accuracy in the free view condition with a mean absolute residual of 5.52 (std=2.84).

5 Conclusion

In this paper, we investigated the link between cognitive health and eye movements during visual activities (e.g., video watching). We first identified a set of eye movement features that correlate to people's memory capability, and then demonstrated that automatic eye movement analysis can predict individual's memory function score (from the standard cognitive tests). These findings provide insights into designing visual tests. Designers can focus on specific eye movement features to assess one's memory health. Our proposed predictive model can potentially be used as a new tool for quantifying cognitive health, without the need of undergoing standard tests at clinics and can be executed in people's home environment. However, our analysis is currently limited because our collected data is from a small sample of subjects in a lab environment. In the future, we intend to evaluate our model on bigger data sets, collected with ambient eye trackers in naturalistic environments. Other than memory capabilities, we will also investigate the relations between eye movements and other cognitive aspects, such as executive function, attention, language skills and more.

6 Acknowledgement

The work described in this paper is funded by EPSRC project EP/M006255/1 Monitoring Of Dementia using Eye Movements (MODEM).

References

1. Anderson, T.J., MacAskill, M.R.: Eye movements in patients with neurodegenerative disorders. Nat Rev Neurol 9(2), 74–85 (Feb 2013)
2. Bednarik, R., Vrzakova, H., Hradis, M.: What do you want to do next: A novel approach for intent prediction in gaze-based interaction. In: Proc. ETRA 2012. pp. 83–90. ETRA '12, ACM, New York, NY, USA (2012)
3. Benson, P.J., Beedie, S.A., Shephard, E., Giegling, I., Rujescu, D., Clair, D.S.: Simple viewing tests can detect eye movement abnormalities that distinguish schizophrenia cases from controls with exceptional accuracy. Biological Psychiatry 72(9), 716 – 724 (2012), cortical Inhibition Deficits in Schizophrenia

4. Borji, A., Itti, L.: Defending yarbus: Eye movements reveal observers' task. Journal of Vision 14(3(29)), 1–22 (Mar 2014)
5. Bulling, A., Roggen, D., Trster, G.: What's in the eyes for context-awareness? Pervasive Computing, IEEE 10(2), 48–57 (April 2011)
6. Crabb, D.P., Smith, N.D., Zhu, H.: What's on tv? detecting age-related neurodegenerative eye disease using eye movement scanpaths. Frontiers in Aging Neuroscience 6(312) (2014)
7. Crawford, T.J., Higham, S., Mayes, J., Dale, M., Shaunak, S., Lekwuwa, G.: The role of working memory and attentional disengagement on inhibitory control: effects of aging and alzheimer's disease. Age 35(5), 1637–1650 (2013)
8. Crawford, T.J., Higham, S., Renvoize, T., Patel, J., Dale, M., Suriya, A., Tetley, S.: Inhibitory control of saccadic eye movements and cognitive impairment in alzheimers disease. Biological psychiatry 57(9), 1052–1060 (2005)
9. Di Stasi, L.L., Renner, R., Staehr, P., Helmert, J.R., Velichkovsky, B.M., Cañas, J.J., Catena, A., Pannasch, S.: Saccadic peak velocity sensitivity to variations in mental workload. Aviation, Space, and Environmental Medicine 81(4), 413–417 (2010)
10. Grober, E., Buschke, H.: Genuine memory deficits in dementia. Developmental neuropsychology 3(1), 13–36 (1987)
11. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Springer Series in Statistics, Springer New York Inc., New York, NY, USA (2001)
12. Henderson, J.M., Shinkareva, S.V., Wang, J., Luke, S.G., Olejarczyk, J.: Predicting cognitive state from eye movements. PLoS ONE 8(5), e64937 (05 2013)
13. Itti, L., Koch, C.: Computational modelling of visual attention. Nature reviews neuroscience 2(3), 194–203 (2001)
14. Jang, Y.M., Lee, S., Mallipeddi, R., Kwak, H.W., Lee, M.: Recognition of human's implicit intention based on an eyeball movement pattern analysis. In: Lu, B.L., Zhang, L., Kwok, J. (eds.) Neural Information Processing, Lecture Notes in Computer Science, vol. 7062, pp. 138–145. Springer Berlin Heidelberg (2011)
15. Jimison, H., Jessey, N., McKanna, J., Zitzelberger, T., Kaye, J.: Monitoring computer interactions to detect early cognitive impairment in elders. In: Distributed Diagnosis and Home Healthcare, 2006. D2H2. 1st Transdisciplinary Conference on. pp. 75–78. IEEE (2006)
16. Kardan, O., Berman, M.G., Yourganov, G., Schmidt, J., Henderson, J.M.: Classifying mental states from eye movements during scene viewing. (2015)
17. Knapp, M., Prince, M., Albanese, E., Banerjee, S., Dhanasiri, S., Fernandez, J., Ferri, C., Snell, T., Stewart, R.: Dementia uk: report to the alzheimer's society. Kings College London and London School of Economics and Political Science (2007)
18. Nasreddine, Z.S., Phillips, N.A., Bdirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., Chertkow, H.: The montreal cognitive assessment, moca: A brief screening tool for mild cognitive impairment. Journal of the American Geriatrics Society 53(4), 695–699 (2005)
19. Schleicher, R., Galley, N., Briest, S., Galley, L.: Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? Ergonomics 51(7), 982–1010 (2008)
20. Steichen, B., Conati, C., Carenini, G.: Inferring visualization task properties, user performance, and user cognitive abilities from eye gaze data. ACM Trans. Interact. Intell. Syst. 4(2), 11:1–11:29 (Jul 2014)
21. Tseng, P.H., Cameron, I., Pari, G., Reynolds, J., Munoz, D., Itti, L.: High-throughput classification of clinical populations from natural viewing eye movements. Journal of Neurology 260(1), 275–284 (2013)