

Screen and Virtual Reality-Based Testing of Contrast Sensitivity

Guillermo Vivas-Mateos
Sarah Boswell
Iain A T Livingstone
Jonathan Delafield-Butt
Mario E Giardini

Accepted at the 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Montreal (Canada) 20-24 July 2020

This is an author-generated copy of the accepted version.

© 2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Screen and Virtual Reality-Based Testing of Contrast Sensitivity

Guillermo Vivas-Mateos *Member, IEEE EMBS*, Sarah Boswell, Iain A.T. Livingstone, Jonathan Delafield-Butt, Mario E. Giardini, *Member, IEEE EMBS*

Abstract— Contrast sensitivity is a key visual ability for everyday tasks, as well as a potential indicator of important optical and neurological diseases. Current clinical standards, based on visual discrimination performance on printed charts, present problems that could be bypassed using electronic devices. This work describes the development of new tests for contrast sensitivity, based on the detection of a moving target on a computer screen and in virtual reality headset. It presents preliminary evaluation of these innovations by comparison of their performance, using healthy adults with normal vision and by artificially altering their contrast sensitivity. The results demonstrate consistent correlation between all test modalities explored.

Clinical Relevance— This study presents new digital tests for the measurement of contrast sensitivity that may offer a pathway to automated testing, reducing the demand on the skillset of the operators.

I. INTRODUCTION

Contrast sensitivity (CS) can be defined as the ability to perceive small differences in luminance between two adjacent targets. It is a primary function of normal vision. In the assessment of visual function, CS testing can be as important as visual acuity testing [1], however, it is not routinely performed [2]. Poor CS can negatively affect the life of the patient, as everyday tasks such as recognizing faces, driving or walking up or down stairs can become challenging [3]. Moreover, low CS can be an indicator of more serious conditions such as glaucoma [4] or optic neuritis [5].

The gold standard for testing CS is the Pelli-Robson (PR) chart, which consists of a set of triplets of letters with diminishing contrast displayed on a white background. The test subject is asked to read the triplets of letters in order of decreasing contrast, until the letters are no longer distinguishable against the background. The last triplet distinguished above a given error rate threshold is recorded as a quantitative score [6]. This test has presented some problems for clinicians such as the large area required for the test, the luminance requirements, and the fact that the test paper is prone to degradation over time, rendering it clinically ineffective after a manufacturer-defined aging period [3]. Another issue with the PR test is that it measures discrimination of a static target. Yet, testing detection of a dynamic target is thought to provide information more representative of the actual tasks encountered during daily life, hence more indicative of the visual ability of the patients [7]. This is because dynamic testing requires a combination of

visual, coordination, and cognitive abilities which match better the use of CS in everyday life such as driving, navigating or competing in sports.

Electronic displays suffer less from aging problems and, in general, can be configured to comply with photometric standards for ophthalmic diagnostics [8]. They have seen a rise in their use for clinical testing, with applications and validation studies for different conditions, including CS, being performed over the past few years [9]–[12]. Further to this, head-mounted virtual reality (VR) is a relatively new technology with great potential in the ophthalmological field because it offers total enclosure of the visual environment, thus allowing full control over the illumination and the stimuli presented [13]. The rise in the use of VR technology for gaming is driving the technology towards higher resolutions, higher response speed, higher availability, and lower prices [14].

In this study, we present the development of both a static and a dynamic CS test, implemented on a computer screen and a VR headset, and we show preliminary evaluation data on a cohort of healthy volunteers with artificially degraded vision.

II. METHODS

Three new tests were developed for CS: an electronic version of the PR chart on a computer screen, a dynamic test displaying a moving dot with decreasing contrast on a computer screen, and a similar moving dot test on a VR headset.

A. *Electronic Pelli-Robson chart*

Our electronic PR chart (ePR) test was developed by following the original PR paper [6], modified to comply with the contrast range implemented in the current version of the chart [3] which is more commonly encountered in clinical practice. In this newer version, the log of Webber contrast values range from 0.00 to 2.25 in 0.15 steps. The ePR was developed using Visual Basic 2017 (Microsoft Inc, USA).

To create the correct luminance values for each optotype triplet, the conventional gray levels offered by the computer screen were not enough. Due to the 8-bit resolution of the color values implemented by the screen / graphics card combination, the gray luminance varied in 256 steps, which are too big to offer a sufficient Webber contrast resolution. Chessboard patterns of 1 pixel square size mixing different levels of gray were therefore created using an OpenCV C++ script [15] to achieve the correct luminance levels (Fig 1.). Such a small size for the chessboard squares made them undistinguishable from

* Research supported by the UK EPSRC (Grant Ref. EP/L015595/1) and Teresa Rosenbaum Golden Charitable Trust (Rosetrees Trust, Grant Ref. M270).

G. Vivas-Mateos, S Boswell, and M.E. Giardini are with the Department of Biomedical Engineering and J. Delafield-Butt with the Laboratory for

Innovation in Autism, all at the University of Strathclyde, Glasgow G1 1XQ, UK (e-mail: g.vivas@strath.ac.uk).

I. A.T. Livingstone is with NHS Forth Valley, Falkirk Community Hospital, Falkirk FK1 5QE, UK (email: iain.livingstone@nhs.net).



Figure 1. The first test, direct translation of paper PR to ePR. Grayscale was derived from the original PR grayscale. In the inset, a small region of the letter S is magnified to show the chessboard pattern used to compose the appropriate shade of gray.

a flat color at the distance of the test, effectively mixing the two levels of gray together [16]. The combinations of gray were chosen through direct measurement of the luminance. The detailed technique goes outside the scope of the present paper.

B. Moving dot test

The moving dot test consisted of a gray dot over a white screen. The dot started in the center of the screen and then moved to both sides alternatively with a horizontal sinusoidal motion on a period of eight seconds (Fig. 2). The test was designed so that the dot started at maximum contrast and was then decreased manually by an operator in real time, using a control panel displayed on a separate screen (a vanishing contrast optotype model).

The moving dot test was developed through two paradigms: (a) a screen-based paradigm and (b) a VR headset-based paradigm. Both paradigms have been used for ophthalmologic applications: screens have been used and tested more; VR offers potential new benefits such as better control over the illumination conditions.

For the screen paradigm, the test was designed in Visual Basic 2017. The gray shades used were emulated using the same chessboard patterns that were used for the ePR (Fig. 1).

The VR-based version of the moving dot test was developed using Unreal Engine 4.19 (Epic Games, USA). Unlike in the screen-based tests, the Unreal Engine environment does not allow for the use of gratings, hence the closest default colors from the RGB color space were used.

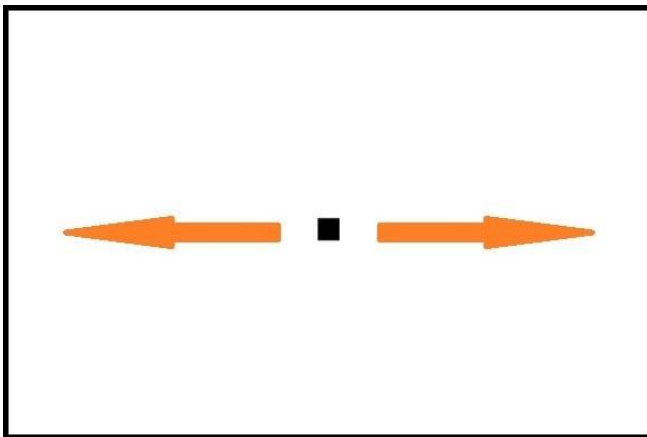


Figure 2. The moving dot test (equivalent for the screen- and VR-based tests). The dot starts in the center of the screen and the moves alternatively to the sides of the screen with a sinusoidal motion.

C. Testing

Nine healthy adult volunteers (three male) with normal vision were recruited from the University of Strathclyde. Inclusion requirements were for them to have corrected acuity of at least 6/9 (checked with a Snellen chart) and no other known visual deficits.

Participants were asked to sit at 1.5 m from a computer screen (Hazro Hz30wiGv1.0, Hazro Technologies) in a black-out room/studio. For the VR-based tests, participants were asked to don a VR headset HTC VIVE (HTC, Taiwan). Each test was administered binocularly and monocularly to each participant. To understand the performance of the tests with different CS levels, the tests were repeated 5 times for each volunteer, namely with normal vision and artificially degrading their vision using Bangerter filters [17] with 4 different strengths: 0.8, 0.4, 0.2, and 0.1 (lower number corresponds to stronger filtering). In total, therefore, the tests were repeated 15 times per participant per technology. The tests were taken in the same order for all participants: 1) ePR 2) screen based moving dot 3) VR-based moving dot.

For the ePR, the participant was asked to read out loud as many letters as possible, considering a triplet as correct when they could read at least 2/3 letters of the triplet. For the moving dot (both screen- and VR-based) the participants were first asked if they could see the static dot in the center of the screen. Following this, the dot was oscillated, and the volunteers were asked to report whether they could see it. The contrast was then reduced, and the test repeated. The test was interrupted and scored when the volunteer could no longer see neither the static nor the moving dot.

D. Statistical analysis

The group samples were small hence the datapoints were not enough to show a normal distribution. Because of this, nonparametric analyses were conducted.

For each filter strength, a Mood's Median test was conducted ($\alpha = 0.05$). Further Mann-Whitney tests were carried out to check which pairs were different where significant differences had arisen from the Mood's Median test.

III. RESULTS

Fig. 3 shows the results from the binocular tests that serve to represent the comparison between the different tests.

A. CS threshold with different filters

All tests returned an inverse relationship between the CS measure and filter strength. At all filter strengths, the screen-based moving dot test returned the highest CS values of all tests. The lowest values were recorded from the ePR or the VR-based static dot. The ePR showed the greatest differentiation between filters, i.e. the greatest changes in CS measured. The VR-based moving dot showed the lowest discrimination between filters.

B. Static vs dynamic testing

The Mood's Median and Mann-Whitney tests revealed that the difference between the dot being static and the dot moving was not statistically significant ($p > 0.05$). This was true for both the screen- and VR-based dot test.

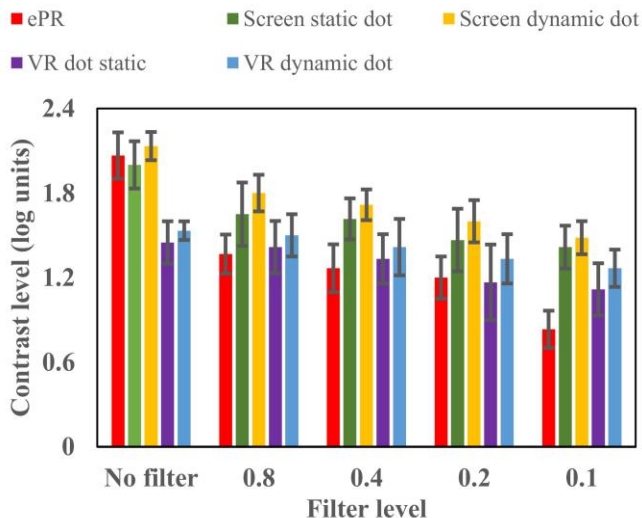


Figure 3. Bar chart of data from binocular results of the testing. The bars represent the average CS threshold and the whiskers the standard deviation of each test for the 5 filter strengths. The results are presented in log units of the Webber contrast, as normally graded in the Pelli-Robson chart. For each filter strength, the bars represent, from left to right: electronic Pelli-Robson (red), screen-based dot test static (green), screen-based dot test dynamic (yellow), VR-based dot test static (purple), VR-based dot test dynamic (blue).

IV. DISCUSSION

Electronic tests present some advantages over the traditional tests: first, the photometric characteristics of the screens may allow for less strict environmental conditions to perform the tests. Electronic devices are usually more accessible and less expensive than specialized tests, as well as more robust to degradation over time. Furthermore, using electronic devices for testing opens the possibility to automate the tests, making them faster and easier to use [18]. All these advantages combined suggest that electronic tests may outperform the more traditional physical tests in current clinical practice and may pose lower demands on the skillset required by the operators, e.g. allowing deployment in community care settings.

The data from this study showed promising results for the performance of the new developed tests. All tests appeared to return different values of CS as the filtering was increased, although to different extents. Odell et al. provide normative values from testing vision artificially degraded by Bangerter filters with a commercial PR chart [17]. For the same filter strengths as used in our study, they obtained the following results (mean \pm std): no filter (1.91 ± 0.12); 0.8 (1.80 ± 0.16); 0.4 (1.65 ± 0.13); 0.2 (1.56 ± 0.14); 0.1 (0.93 ± 0.12). Our ePR test showed similar results for the no filter and 0.1 filter but were very different in the rest. Regarding the screen-based dot test (both static and dynamic) the test returned similar values for all filter strengths except the 0.1. Lastly, the VR-based dot test did not match the results for any filter strength.

Considering these comparisons, as well as the sensitivity to the changes in CS produced by the Bangerter filters, the screen-based dot test appears to offer the best match, while the VR-based dot test appears to offer the worst. The ePR matched better the existing data for very low CS levels (Bangerter filter 0.1). While we report this as an objective consideration, we are somewhat cautious in drawing conclusions from this, as

Bangerter filters are designed for eye occlusion, a significantly different purpose from our work and, in terms of contrast degradation, are anecdotally known in the community of practice not to offer reliable repeatability.

Two factors are thought to cause the seemingly worse performance of the VR-based test: first, the inability to use chessboard patterns in the VR environment, which may have caused luminance mismatches with respect to the theory. Second, the screen-lens system from the HTC headset conspicuously appears to have a much larger chromatic aberration than common LCD screens, and clearly visible pixel size. At this point, VR technology does not appear to be mature enough to hold these tests.

Although there were no statistical differences between the CS threshold measured by detecting the dot being static or moving, at an individual level, all participants had an equal or higher CS score when the dot was moving rather than when it was static. The differences recorded ranged between zero and two contrast levels (0 to 0.30 log). The reason for this is unclear at this stage, it is theorized it could be because of the difference between static and dynamic visual acuity [7], flicker sensitivity [19] or indeed relate to the different way motion information is processed at the retina, lateral geniculate nucleus and visual cortex.

There are three factors that we think could enhance the results from this study in further development stages: first, trialing the tests with actual patients with low CS instead of artificially altering it via Bangerter filters. Although the filters do reduce CS, it has also been proven that they reduce near, distance and Vernier acuity [17], which could be affecting the results obtained. Second, including a current commercial PR chart would improve the comparison as it would compare results from the same group of participants rather than against normative data from other studies. Last, increasing the sample size of the study would also increase the robustness and veracity of the analysis.

The letter chart presents a significant conceptual difference from the dot test. Indeed, while reading letters tests CS with a discrimination task, the dot tests target detection. Discrimination presents some advantages over detection: the letter chart test is faster than the moving dot test since reading letters is faster than observing the dot for several seconds, several times, with different contrast. Another advantage is that the letter chart is less prone to false positives: if the letters are named correctly, there is a high chance that the patient is factually discriminating them. On the contrary, the normal noise in the visual process can easily be confused with a dot at low contrast, be it static or in movement, especially if there is an expectation on the position of the dot. This second point could be solved by using head or gaze tracking technologies to track the pursuit of a dot following a complex trajectory on the screen. With respect to the letter chart, the dynamic dot paradigm has more potential to be automatized if head and gaze technologies are available, which could also open it to a wider public such as infants, illiterate people, and adults with special needs, where literacy cannot be assumed. Similarly, the dynamic dot paradigm presents advantages regards global usefulness as it is independent regards language.

V. CONCLUSION

New electronic screen and VR-based tests were developed to measure CS and tested with healthy volunteers with artificially degraded vision. Although the VR-based tests did not return encouraging results, the results obtained from the screen-based tests give cause for further studies to understand the performance and clinical significance on actual patients suffering from low CS.

APPENDIX

This study was approved by the research ethics committee of the Department of Biomedical Engineering of the University of Strathclyde in compliance with the Declaration of Helsinki and UK law.

REFERENCES

- [1] G. B. Arden, "The importance of measuring contrast sensitivity in cases of visual disturbance," *Br. J. Ophthalmol.*, vol. 62, no. 4, pp. 198–209, 1978.
- [2] M. Dorr, T. Elze, H. Wang, Z. L. Lu, P. J. Bex, and L. A. Lesmes, "New precision metrics for contrast sensitivity testing," *IEEE J. Biomed. Heal. Informatics*, vol. 22, no. 3, pp. 919–925, May 2018.
- [3] A. Arditi, "Improving the design of the letter contrast sensitivity test," *Investig. Ophthalmol. Vis. Sci.*, vol. 46, no. 6, pp. 2225–2229, 2005.
- [4] J. E. Ross, "Clinical detection of abnormalities in central vision in chronic simple glaucoma using contrast sensitivity," *Int. Ophthalmol.*, vol. 8, no. 3, pp. 167–177, 1985.
- [5] J. D. Trobe, R. W. Beck, P. S. Moke, and P. A. Cleary, "Contrast sensitivity and other vision tests in the optic neuritis treatment trial," *Am. J. Ophthalmol.*, vol. 121, no. 5, pp. 547–553, 1996.
- [6] J. G. R. A. J. W. J. D. G. Pelli, "The design of a new letter chart for measuring contrast sensitivity," *Clin. Vis. Sci.*, vol. 2, no. 3, pp. 187–199, 1988.
- [7] G. M. Long and M. J. Zavod, "Contrast sensitivity in a dynamic environment: Effects of target conditions and visual impairment," *Hum. Factors*, vol. 44, no. 1, pp. 120–132, 2002.
- [8] I. A. T. Livingstone, C. M. Tarbert, M. E. Giardini, A. Bastawrous, D. Middleton, and R. Hamilton, "Photometric compliance of tablet screens and retro-illuminated acuity charts as visual acuity measurement devices," *PLoS One*, vol. 11, no. 3, pp. 1–12, 2016.
- [9] T. M. Aslam, I. J. Murray, M. Y. T. Lai, E. Linton, H. J. Tahir, and N. R. A. Parry, "An assessment of a modern touch-screen tablet computer with reference to core physical characteristics necessary for clinical vision testing," *J. R. Soc. Interface*, vol. 10, no. 84, 2013.
- [10] H. K. Rono *et al.*, "Smartphone-based screening for visual impairment in Kenyan school children: a cluster randomised controlled trial," *Lancet Glob. Heal.*, vol. 6, no. 8, pp. e924–e932, 2018.
- [11] P. S. Kollbaum, M. E. Jansen, E. J. Kollbaum, and M. A. Bullimore, "Validation of an iPad test of letter contrast sensitivity," *Optom. Vis. Sci.*, vol. 91, no. 3, pp. 291–296, Mar. 2014.
- [12] M. Dorr, L. A. Lesmes, Z. L. Lu, and P. J. Bex, "Rapid and reliable assessment of the contrast sensitivity function on an iPad," *Investig. Ophthalmol. Vis. Sci.*, vol. 54, no. 12, pp. 7266–7273, Oct. 2013.
- [13] N. Nesaratnam, P. Thomas, and A. Vivian, "Stepping into the virtual unknown: Feasibility study of a virtual reality-based test of ocular misalignment," *Eye*, vol. 31, no. 10, pp. 1503–1506, 2017.
- [14] D. C. Niehorster, L. Li, and M. Lappe, "The accuracy and precision of position and orientation tracking in the HTC vive virtual reality system for scientific research," *Iperception.*, vol. 8, no. 3, pp. 1–23, 2017.
- [15] "OpenCV." [Online]. Available: <https://opencv.org/>. [Accessed: 20-Jan-2020].
- [16] I. A. T. Livingstone, A. S. L. L. Lok, and C. Tarbert, "New mobile technologies and visual acuity," *2014 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBC 2014*, pp. 2189–2192, 2014.
- [17] N. V. Odell, D. A. Leske, S. R. Hatt, W. E. Adams, and J. M. Holmes, "The effect of Bangerter filters on optotype acuity, Vernier acuity, and contrast sensitivity," *J. AAPOS*, vol. 12, no. 6, pp. 555–559, Dec. 2008.
- [18] R. Clark *et al.*, "The potential and value of objective eye tracking in the ophthalmology clinic," *Eye (Basingstoke)*, vol. 33, no. 8. Nature Publishing Group, pp. 1200–1202, 01-Aug-2019.
- [19] A. R. Seitz, J. E. Nanez, S. R. Holloway, and T. Watanabe, "Perceptual learning of motion leads to faster flicker perception," *PLoS One*, vol. 1, no. 1, pp. 1–9, 2006.