Towards a remote EEG for use in robotic sensors

Niels-Ole Rohweder[1], Christian Rembe[1], and Jan Gertheiss[2]

[1] Clausthal University of Technology,
Institute for Electrical Information Technology,
Leibnizstrasse 28, 38678 Clausthal-Zellerfeld
[2] Helmut Schmidt University, Department of Mathematics and Statistics,
Holstenhofweg 85, 22008 Hamburg

Abstract Recently, it was shown that a correlation exists between brain activity and oscillations of the pupil. As the experiment was designed to measure exitations of the pupil for frequencies below 1 Hz, whether such correlations also exist on the scales of seconds and for frequencies between 5 and 40 Hz is still an open question. In this work, we design a new experiment and measure the response of the pupil to continuous, periodic visual and acoustic stimuli. We show that a clear response of the pupil for flashes of 7.5 Hz exists, bearing similarity to the effect known as Steady-State Visual Evoked Potential in neuroscience. This result can directly be used to develop a new kind of non-contact brain-computer-interface, using visual fixation as a trigger. Further, we evaluate the pupil response to series of white noise clicks with a frequency of 8 Hz, in order to assert the pupil response as due to brain activity. First results indicate the presence of a weak signal, showing the stimulus frequency and harmonics, bearing similarity to the neural effect known as Auditory Steady-State Response. Measuring brain activity remotely could provide pathways to new kinds of sensors, in particular for collaborative robots and general human-machine-teams, where estimates of the mental state of the human partner are essential.

Keywords eeg, remote eeg, pupil, oscillations, SSVEP, ASSR, HMT, BCI, sensor

DOI: 10.58895/ksp/1000124383-18 erschienen in:

Forum Bildverarbeitung 2020

DOI: 10.5445/KSP/1000124383 | https://www.ksp.kit.edu/site/books/m/10.58895/ksp/1000124383/

1 Introduction

One central aspect of Human-Machine-Teams (HMT) – be it collaborative work in factories, or driving partly autonomous cars – is the ability of the machine to evaluate its human partner. Especially in safety-critical environments, a precise state model of the human is essential, consisting, at a minimum, of a binary attribute of "take-over-readiness" (TOR) – the ability of the human to perform the required task, e. g. taking back control of the steering wheel, or accepting a hazardous object in a factory. Otherwise, the machine or robot is left blind; and indeed, that is the current state-of-the-art. In collaborative situations such as autonomous cars, the focus is placed on clear interfaces to signal the human partner the need to take over, without making sure they are actually able to do so [1]. Even the very term of "take-over-readiness" and the concept it describes is used solely in the context of partly autonomous cars, not found anywhere else, and the application of human models and considerations of the HMT as one unit, i. e. a human-in-the-loop approach, are only of very recent focus in the literature [2, 3].

Regardless of the complexity of the human model – a single attribute or a full Theory of Mind – the required sensor input is going to consist of both physical and mental parameters. While research and sensors for both exist, only physical parameters (e. g. hand position, heart rate) so far have been measured remotely. The gold standard for measuring mental parameters such as situation awareness, cognitive load or tiredness, the electroencephalogram (EEG), requires electrodes placed on the scalp for good signal quality. Obviously, such a setup is infeasible in real world applications, outside of very special circumstances. On the other hand, Park and Whang recently showed that a correlation exists between brain activity and pupil size, such that the electrical state changes created by the neurons – and measured as oscillations of electrical potential on the skin – are mirrored by oscillations of the size of the pupil [4]. In particular, they showed a strong correlation of activity in the front and central brain region with pupil oscillations, e. g. in the mu-band around 10 Hz, and the gamma-band between 30 and 50 Hz.

However, the setup of Park and Whang used long-time averages, comparing the frequency bands of the EEG with 1/100 subharmon-

ics for the pupil oscillations (e. g. the 10-Hz-band of the former with pupil oscillations around 0.1 Hz), thus creating correlations on the scales of minutes. A natural extension of the experiment is to ask whether such correlations also can be measured directly, using the same frequency bands, thus increasing the time resolution to seconds, rather than minutes. A second question is whether such correlations, if they exist, can be used to create a reliable remote EEG, for use in sensors to serve as input of a human model, e. g. in the context of deriving a measure of "trust" between the partners [5]. This overachieving question is the purpose of the HerMes project of the Clausthal University of Technology, in the context of which our experiments are performed [6].

2 Steady-State Evoked Potentials as stimuli in experiments

From the outset, it is clear that measuring oscillations of frequencies higher than $\sim$1 Hz decreases the resulting amplitude drastically. The pupil is an imperfect oscillator, the speed of the dilation or contraction which the iris muscles can achieve is limited, hence the response to any stimuli is limited as well. Literature confirms this hypothesis [7, 8]. In order to increase the signal-to-noise ratio, it is therefore desirable to have an artificially triggered, continuous signal that can clearly be separated from the noise and measured for arbitrary durations. In neuroscience, such signals are known as "Steady-State Evoked Potentials", a response of the brain to a continuous sensory stimulus at a certain frequency. The stimulus frequency, typically including harmonics, can be clearly measured in the brain activity, for as long as the stimulus persists [9].

The visual version – the Steady-State Visual Evoked Potential, SSVEP – is the one most easily measured. It has an excellent signal-to-noise ratio (SNR) [10]. It is easily triggered e. g. by flashing LEDs or flickering computer screens, and often used for brain-computer-interfaces [11]. The acoustic variant – Auditory Steady-State Response, ASSR – has at least an order of magnitude smaller responses compared to the SSVEP, and is comparatively harder to measure, typically averaged over multiple traces or long periods to create a

sufficient SNR. Generally, the strongest responses appear at and below frequencies of 40 Hz [12]. Both visual and acoustic stimuli were identified as candidates to measure a response in the pupil oscillations.

3 Algorithm

In order to resolve such small pupil oscillations, a precise computer algorithm for evaluating the pupil diameter is needed. Typically, a circle detection is used. Yet unless the camera is placed such that it is perpendicular to the pupil, this is an approximation; the perspective distortion turns the circle into an ellipse of increasing eccentricity for increasing angles. In experiments such as Park and Whang's, as well as in commercial eye-trackers such as Tobii [13], this "pupil foreshortening error" is usually ignored; or avoided, by using large distances between eye and camera. Protocols for post-hoc correction exist [14]. On the other hand, the most convenient placement of the camera for high resolutions of the pupil is close to it, and out of the line of sight, i. e. nearly always at an angle, e. g. looking up from below (Fig. 3.1).

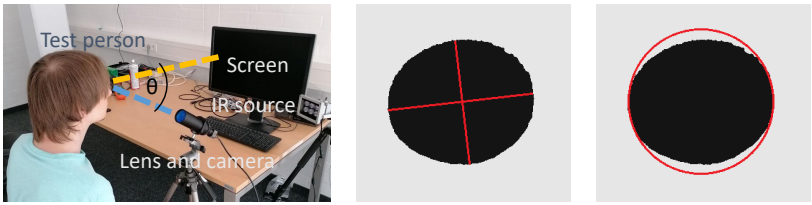

Figure 3.1: Left: The setup of our experiment. Centre: After thresholding to binarise the image of the eye, major and minor axes of the black pupil are detected. The aspect ratio in this image is 1.17, corresponding to an angle θ of approx. 31° for the camera inclination. Right: The perspective distortion causes the pupil to deviate from the ideal circle.

Thus the experiment was planned to improve the previous setups, using elliptical pupil tracking from the start. However, while increasing the accuracy, it also increases complexity. A circle detection algorithm, e. g. the commonly used Hough-transform, operates in a

three-dimensional parameter space, corresponding to the three parameters defining the circle: The x and y component of the centre, as well as the radius. Conversely, an ellipse creates a five-dimensional parameter space, introducing two radii (or axes) and a rotation angle, in addition to the centre coordinates. In order to calculate the relevant quantity, the long (semi-)axis, we use an approach derived from [15, 16]. From the picture to the result, the algorithm works as follows.

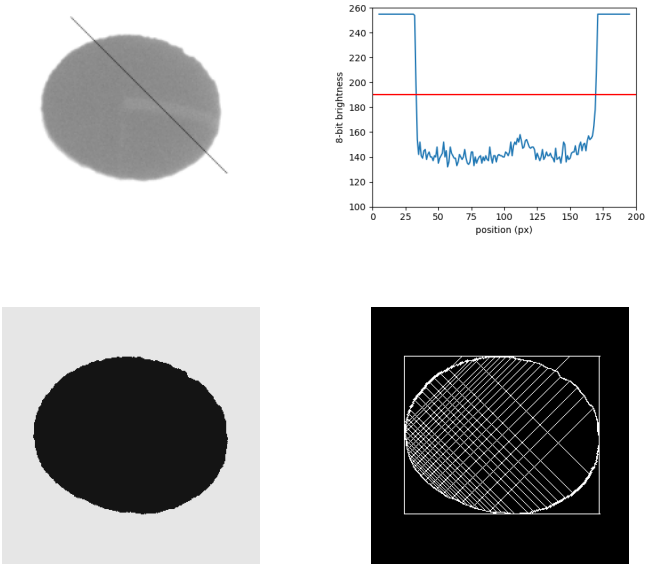

Figure 3.2: Steps of the algorithm. Top left: Image of pupil. A reflection of the experiment's screen is faintly visible, the line indicates a cross-section. Top right: Cross-section along the line with binarising threshold in red. Bottom left: Binarised image. Bottom right: Final result of the Canny edge detection, with the chords used to calculate the ellipse centre shown for illustrative purposes. Chords exceeding the box of the ellipse such as near the upper right corner, due to e. g. missing edge pixels, are discarded.

As the setup is such that the greyscale-image of eye results in the pupil being darker than the rest of the image (see Sec. 4), it is rather easy to binarise, leaving only the pupil behind (Fig. 3.2). Over the binarised image runs a conventional Canny algorithm, which filters

out the edge pixels. The interior is then divided by two sets of chords, running from one side of the ellipse to the other, starting and ending at the innermost pixels. The number of the chords scales with the size of the ellipse, at a typical size of around 20 for each set. Start and end points of the chords create a set M of edge points, defining the ellipse. In accordance with [15,16], we use a geometrical property to determine the centre of the ellipse: The bisections of two sets of chords intersect at the centre of the ellipse. Using a linear regression over the individual centre pixels of a set of chords provides an accurate estimate of its bisector, and leads to a very good estimate of the ellipse centre.

After determining the centre coordinates, the remaining parameter space is three-dimensional once more, and the other parameters can be solved iteratively and algebraically, using three points of the set M. To that end, M is sorted by quadrants and the three points taken from three different quadrants. These three points need to be far from respective points of symmetry that would result in an indetermined ellipse, such as would be the case e. g. with two points $(X|Y), (-X|-Y)$ as measured from the ellipse centre. The average of the result from each of the sets of three is finally taken to get the desired result, the value of the long and the short axis.

The aim here is real-time capability; currently, the entire algorithm, from saving the image to getting the axes lengths, runs at approximately 20 fps on an Intel core i5 (8th gen) processor.

4 Experimental setup

The three parts of the experiment are the test person, the camera and the source of light. The test person takes a seat in front the computer screen. The eye is filmed from below, while the look fixes ahead, on a mark on the screen. The illumination comes from the side (Fig. 3.1).

There are two ways to create a high contrast between pupil and the rest of the eye or face. For the bright pupil effect, in which light is reflected back from the retina into the camera (the same as the "red-eye effect" in photography with a flash), camera and source of light are required to be closely aligned on one optical axis. This in turn

requires a sufficient distance of the eye from the camera, typically on the order of metres, as the dimensions of the camera and the source of light limit their proximity. The dark pupil effect does the inverse, placing the source of light off-angle; the light is reflected away from the camera, and the pupil appears black. This is the only feasible option, as the camera is placed less than half a metre away from the eye in order to create a sufficient resolution of the pupil.

As usual to avoid glare and to provide uniform illumination, near infrared light is used. A simple 6 W LED floodlight of 850 nm wavelength is placed such that the specular reflection off the cornea is not directed towards the camera, avoiding a bright glint in the otherwise dark pupil. The lens in front of the camera is assembled using two Near-IR coated lenses of $f = 150$ and $f = 25.4$ mm focal length and an iris diaphragm in their common focal point, creating a simple telecentric lens. The advantage of such a setup is the independence of the magnification from the distance of the object.

On the one hand, this avoids changes in pupil size by involuntary movements of the head, and on the other hand allows to calibrate the optical system such that the pupil oscillations can be examined in real units, not pixels. As the amplitude of such oscillations has never been measured and therefore is unknown, this was an important consideration. The calibration was performed before the experiment using an USAF target, and determined as 20.83 ± 0.12 µm per pixel. The lens is mounted on a camera with a 3.2 MP resolution (IDS UI-3270CP), of which a field of view of 1024 by 1024 pixels is used, allowing for frame rates of 83 fps. Consequently, frequencies up to 40 Hz can be resolved.

5 Results and discussion

For the first part of the experiment, an SSVEP sequence is displayed on the screen. It consists of 16 seconds of black screen, then 10 seconds of a white flashing box, and an additional 10 seconds of black screen at the end. The frequencies of the flashes are chosen such that the refresh rate of the screen can be synced, i. e. 30 Hz, 15 Hz, 10 Hz or 7.5 Hz. An exemplary result is displayed in Fig. 5.1 for a stimulation frequency of 7.5 Hz. Both the SSVEP stretch and the

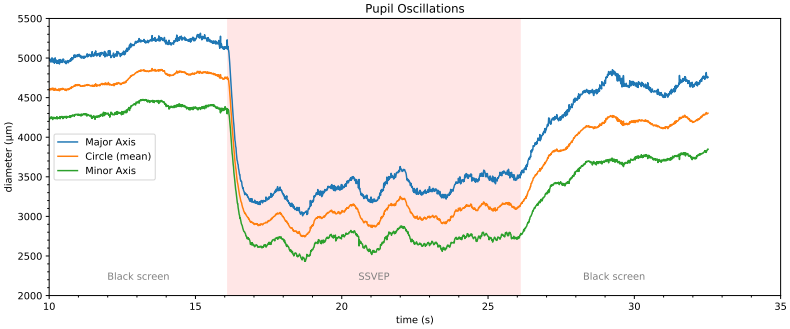

Figure 5.1: Pupil diameter before, during, and after a sequence. The SSVEP stretch is highlighted.

dark screen can be clearly determined by the pupil size. The relative brighter stretch of the flickering screen causes the pupil to contract. A Fourier analysis (resolution bandwidth: 0.125 Hz) of the SSVEP-stretch yields a clear peak at the stimulation frequency of 7.5 Hz, as well as one at its third harmonic, 22.5 Hz (Fig. 5.2, left). The signal-to-noise ratio of the fundamental was 17 dB. The amplitude of the oscillations is on the order of 10 µm. As one pixel was calibrated to 20.83 µm, this is a sub-pixel resolution, a result of the parameter estimation algorithm and the Fourier Transform over a sufficiently long time interval.

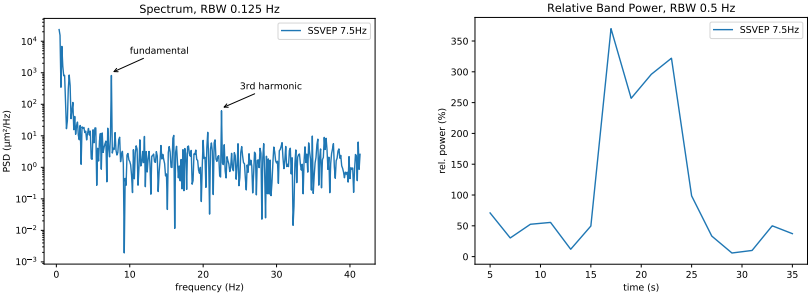

Figure 5.2: Left: Power spectral density of the recorded SSVEP sequence with a frequency of 7.5 Hz. Right: Relative power of the 7.5 Hz band over time, with respect to the mean power of four bands (7.5 Hz, 10 Hz, 15 Hz, 30 Hz).

By calculating the mean power of the four bands (30 Hz, 15 Hz, 10 Hz and 7.5 Hz, resolution bandwidth 0.5 Hz), and looking at the relative power of the respective individual bands with regards to that mean, Fourier transforming intervals as short as 2 seconds resulted in observable pupil responses (Fig. 5.2, right). This result has direct relevance for creating new kinds of non-contact brain-computer-interfaces. By focusing on one of multiple flickering spots, each with a different frequency, and using a threshold value on such a time series of band power, the resulting potential can trigger actions, e. g. controlling disability aids such as wheelchairs [11]. The idea is similar to [7], who suggested such a scheme, at lower frequencies, for tracking attention or focus.

Unfortunately, for visible stimuli, it is hard to separate oscillation induced via the steady-state potential and brain activity from oscillations due to the simple pupillary light reflex. The amplitude of either oscillation is limited by the mechanical constraints of the iris muscle at any rate; biological considerations such frequency-dependent light reflex responses and latencies to estimate the influence on the measurements would appear to increase the complexity of the experiment severely. Instead, we chose to investigate the response of the pupil to acoustic stimuli. This creates a clear distinction between light reflex and brain activity-induced oscillations; however, as noted earlier, the ASSR effect is at least an order of magnitude smaller than its optical counterpart, and therefore harder to isolate from the noise.

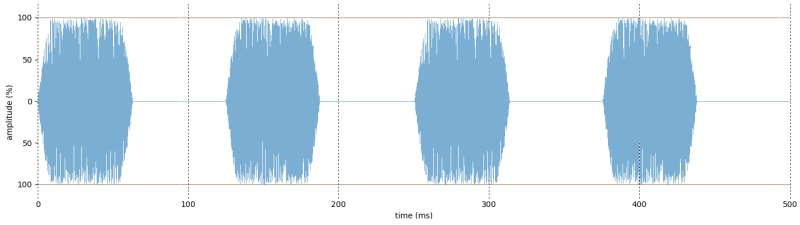

Figure 5.3: The ASSR stimulus. White noise, amplitude-modulated with a rectangle wave of 8 Hz. The modulation depth is 100%.

We used series of white noise click trains, that is, a 100% amplitude modulated sequence of white noise lasting ten seconds and incorporating 80 clicks, creating a rectangle wave of 8 Hz (Fig. 5.3).

The soundfile is played using headphones, while the look of the test person is fixed ahead onto a permanent mark on the screen. The rest of the procedure, as well as the evaluation of the recorded images are exactly as above, using Fourier transforms with a resolution bandwidth of 0.125 Hz.

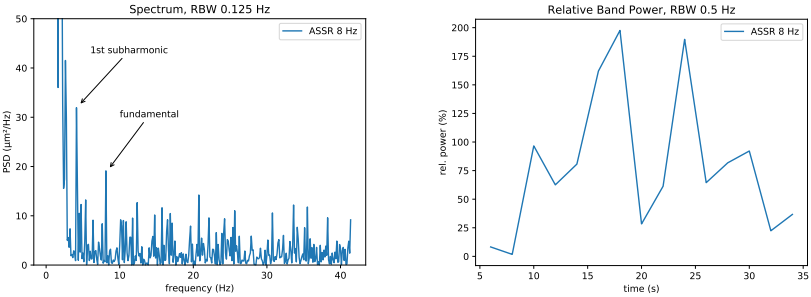

Figure 5.4: Left: Power spectral density of the recorded ASSR sequence with a frequency of 8 Hz. Right: Relative power of the 8 Hz band over time, with respect to the mean power of four bands (8 Hz, 10 Hz, 15 Hz, 30 Hz).

First results indicate the presence of the AM stimulus frequencies and (sub-)harmonics (Fig. 5.4, left); however, the recordings are subject to a lot of noise, often entirely covering the signal. The time series of the relative band power (Fig. 5.4, right) indicates the this: The sequence starts after 15 seconds, and for the 10 second duration, the signal is sporadic, fading entirely before reappearing. The signal-to-noise ratio never exceeds 3 dB, the measured amplitudes of those oscillations are below 4 µm. Interestingly, for highly eccentric ellipses, the oscillations, if they do appear, are visible in the diameter of the long axis only, not in the short axis, likely because the squeezed amplitude due to the perspective distortion is too small to resolve. This justifies the use of the elliptical fit.

Nevertheless, further improvements to both the algorithm and the setup may be needed, to increase resolution and precision and decrease noise, and achieve robust results. A commercial EEG to measure brain activity directly, which so far has not been used, will create another way to confirm the correlation.

6 Summary

A remote EEG is a promising way to create a sensor for use in Human-Machine-Teams. Building on the work of Park and Whang, we developed an experiment to show a correlation between brain activity and oscillations of the pupil diameter. As opposed to Park and Whang, we tried measuring the oscillations in the same frequency band as the brain activity, not its subharmonics, as well as in real units, not pixels. In order to achieve the required precision, we placed the camera in close proximity to the eye for a high resolution of the pupil, and developed an elliptical fitting algorithm in order to compensate perspective distortion. Using both visual and acoustic stimuli – SSVEP and ASSR – we tried to measure the corresponding excitation of the pupil. For the visual stimulation, we received a clear and reliable signal; oscillations of the pupil of approximately 10 µm for a stimulation frequency of 7.5 Hz with a SNR of 17 dB. Using intervals for the Fourier transform as short as 2 seconds, we created a time series of the band power, showing the onset as well as the end of the stimulation, thus allowing for the construction of non-contact brain-computer interfaces. In order to separate the brain activity-induced oscillations from the pupillary light reflex, we also used acoustic stimuli. First results indicate a positive response, showing the stimulation frequency of 8 Hz as well as their (sub-)harmonics; however, subject to much noise, sometimes blanketing the signal, and never exceeding an SNR of 3 dB. The corresponding amplitude of the oscillation is below 4 µm. In the future, decreasing the noise floor and correlating the pupil spectra with commercial EEG spectra could yield more robust results.

References

1. C. Braunagel, W. Rosenstiel, and E. Kasneci, "Ready for take-over? A new driver assistance system for an automated classification of driver take-over readiness," *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 4, pp. 10–22, 2017.
2. N. Deo and M. M. Trivedi, "Looking at the Driver/Rider in Autonomous Vehicles to Predict Take-Over Readiness," *IEEE Transactions on Intelligent*

Vehicles, vol. 5, no. 1, pp. 41–52, 2019, conference Name: IEEE Transactions on Intelligent Vehicles.

3. T. Mioch, L. Kroon, and M. Neerincx, "Driver readiness model for regulating the transfer from automation to human control," in *Proceedings of the 22nd international conference on intelligent user interfaces*, Mar. 2017, pp. 205–213.
4. S. Park and M. Whang, "Infrared camera-based non-contact measurement of brain activity from pupillary rhythms," *Frontiers in physiology*, vol. 9, 2018.
5. B. Alhaji, J. Beecken, R. Ehlers, J. Gertheiss, F. Merz, J. Müller, M. Prilla, A. Rausch, A. Reinhardt, D. Reinhardt, C. Rembe, N. Rohweder, C. Schwindt, S. Westphal, and J. Zimmermann, "Engineering human-machine teams for trusted collaboration," *submitted to: Big Data and Cognitive Computing*, 2020.
6. HerMes, https://www.simzentrum.de/en/hermes/.
7. M. Naber, G. A. Alvarez, and K. Nakayama, "Tracking the allocation of attention using human pupillary oscillations," *Frontiers in Psychology*, vol. 4, 2013.
8. P. A. Barrionuevo, N. Nicandro, J. J. McAnany, A. J. Zele, P. Gamlin, and D. Cao, "Assessing Rod, Cone, and Melanopsin Contributions to Human Pupil Flicker Responses," *Investigative Ophthalmology & Visual Science*, vol. 55, no. 2, pp. 719–727, Feb. 2014.
9. C. S. Herrmann, "Human EEG responses to 1-100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena," *Experimental Brain Research*, vol. 137, no. 3-4, pp. 346–353, Apr. 2001.
10. D. Regan, *Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine*. New York: Elsevier, 1989.
11. B. Allison, I. Sugiarto, B. Graimann, and A. Gräser, "Display optimization in SSVEP BCIs," in *Computer-Human Interaction*, 2008, pp. 2–5.
12. T. W. Picton, M. S. John, A. Dimitrijevic, and D. Purcell, "Human auditory steady-state responses: Respuestas auditivas de estado estable en humanos," *International Journal of Audiology*, vol. 42, no. 4, pp. 177–219, Jan. 2003.
13. Tobii, https://www.tobii.com/.
14. T. R. Hayes and A. A. Petrov, "Mapping and correcting the influence of gaze position on pupil size measurements," *Behavior research methods*, vol. 48, no. 2, pp. 510–527, Jun. 2016.

15. S.-C. Zhang and Z.-Q. Liu, "A robust, real-time ellipse detector," *Pattern Recognition*, vol. 38, no. 2, pp. 273–287, Feb. 2005.

16. A. M. Fernandes, "Discussion on paper "A Robust Real-Time Ellipse Detector" by Zhang and Liu," *Pattern Recognition*, vol. 44, no. 2, pp. 488–489, Feb. 2011.