The Receiving End of the Cooperative Eye Hypothesis

M Schmettow, C. Willemse, S. Borsci

## Introduction

The human eye region has a central role in non-verbal communication. Next to emotional expressions, humans are very accurate and fast at reading each others glance directions. The Cooperative Eye Hypothesis (CEH) states that this ability has evolved, when human ancestors developed their ability to cooperate and show genuine altruistic behaviour. Initially, the hypothesis is based on the observation that Homo Sapiens (HS) is the only species among extant Great Apes, which has developed a visible white sclera. The first argument in this paper is that under the CEH it is very likely that the eye ball evolves into a form that is *computationally efficient*. The ability to establish shared attention by glance perception may have also played a role in child rearing, which suggests that a *specialized receiver mechanism* to read glance direction may have evolved alongside.

Eye tracking devices are a instruments to observe glance directions. From the CEH it follows that the eye ball is sending strong signals, such that a device to read glance directions should be easy to construct, if just the cognitive mechanism was known. Under the CEH, effective eye tracking must be possible using only visual information that also is available to a human receiver. In contrast, recent eye tracking technology often relies on corneal reflections, which is not what humans usually see and thus not cannot be biomimetic.

Truly bio-mimetic eye tracking devices would require a precise cognitive model of glance perception, which is not available, yet. However, it is possible to build an eye tracker, that can be called *bioconvergent*, in that it may work differently in detail, but limited to the information available also to human receivers. Moreover, given the exceptional performance of human glance perception, a bio-convergent device, must be accurate and computational efficient. In this paper we present a bio-convergent method for eye-tracking. The *QuadBright* method uses visible light brightness distributions of the image and basic statistical learning to establish eye tracking. The first study presented in this paper evaluates the accuracy of the QuadBright method.

Once a bio-convergent eye tracking device is established, it is the first candidate for a cognitive model of human glance perception. This can be tested experimentally, by comparing human glance perception performance with the information available to the QuadBright algorithm, to human glance perception accuracy with the original image. Human glance perception is known to be extremely quick and effortless under normal conditions, a follow-up test for a candidate model is how it performs under brief exposure times.

The second study presented in this paper evaluates human glance perception performance in a two-factorial experiment, comparing the accuracy of the Quadbright degradation to the original image under varying exposure times.

The aim of this paper is to first test the CEH by construction. If the Quadbright algorithm is indeed effective for (machine-based) eye tracking, the claim that the eyeball has evolved to be a computationally efficient sender of glance direction signals is supported. The second aim is to extend the CEH by postulating that a reciprocal cognitive mechanism has evolved on the receiving end, and proposing a candidate model for it.

### The Cooperative Eye Hypothesis (CEH)

All animal scleras are white, but in almost all species most oof it is occluded by a layer of pigmentation in the conjunctiva. As KK note, sustaining pigmentation costs the organism energy. From a biophysical perspective it takes a very precise molecular configuration to produce exact colors. The biosynthetic pathways of melanin are complex and several mutations are known, such as albinism, and bright iris color (Eiberg et al., 2008). Non fatal single-point mutations are possible and could have benefits for the organism, the real question therefore seems to be: why have Great Apes and many other animals dark scleras?

One explanation is that protection against predators. Since the eyeball is an optical device in the first place, a pupil is dark almost inevitable. In its default state, the animal eye is a high-contrast target, that is also rapidly moving in the case of Great Apes, and the conjunctiva pigmentation may have evolved as camouflage (KK). Another hypothesis is that within-species competition is the driving force for dark scleras in the Great Apes.

### Origins of Glance Reading

Originally, the Cooperative Eye Hypothesis (CEH) by KK is based on one distinguishing, and literally eye-catching, feature of Homo Sapiens anatomy. However, there is no signal without a receiver. For the CEH to be complete, it must explain how the receiving mechanism has evolved, unless the human eyeball has evolved to such a perfection that it entirely works on existing cognitive mechanisms.

Tracking down the evolutionary origins of cognitive processes, in this case, reading glances, is much more intricate. The best we know comes from a comparative studies with chimpanzees [REF], where the participants showed no sign of glance reading, but rather relied on head position. If only we assume that chimpanzees have not lost the ability to read glances since the last common ancestor with HS, the ability to read glances must have evolved in the last two million years.

If pigmented eyes have developed for camouflage, the possibility of a single-point mutation or a genetic drift producing a white sclera certainly exists. However, even if the human eyeball sprang into existence by a single-point mutation, evolutionary theory requires that it must have had at least some immediate reproductive advantage. For example, it could suffice that reading white-sclera glances is an ability that can be learned, rather than being innate. A possible way to test this hypothesis is to train chimpanzees to read glance directions.

The remaining question therefore is whether the white sclera of modern humans appeared as a single-point mutation, or by slow genetic drift. Assuming that the receiver mechanism at that time was effective only to a fraction of the current human glance reading performance, the white sclera must have been an effective signal from the start, making a single-point mutation more likely. Perhaps, this was followed by a slow genetic drift on the side of the sender.

### Cognitive mechanisms of glance reading

Face detection is probably the first specialized recognition for complex visual pattern to emerge in human ontogenesis. It is known to be extremely fast [REF] and robust, even greedy, as is evident by the phenomenon of anthropomorphism [REF]. Face detection is doubtlessly a cognitive function shaped by evolution, but there are differences to glance direction reading.

First of all, almost all vertibrates have faces and several mammals from different branches use facial expressions in social interaction [REF] (felidae, canae, apes). Even if this would be a result of convergent evolution, we assume that face detection has evolved much longer ago than glance direction reading.

From a perceptual perspective, a face is a configuration of low-detail “visual blobs”. As the human visual system detects low-frequency features much faster, than fine details, it is easy to see how face detection could become such a feat. However, in the usual range of social distance of 1-2m, the eye region consumes a much smaller visual angle, and the transmitted information is much more delicate, as small movements of the eye ball can change the glance direction by several degrees.

Yorzinski et al. (2021, 10.3389/fpsyg.2021.632616) evaluated the glance reading performance in various conditions, including pigmentation of sclera and iris color. The found a strong increase in latency in the combination of pigmented sclera with a dark iris. More interestingly for the quest of the receiving mechanism is that the accuracy was only affected very slightly. Overall, this experiment shows that glance reading is robust under a variety of conditions.

…

As established above, some cognitive structures must have existed on the receiving end, when the white sclera appeared. On the lowest level of visual processing, the human visual system is attracted by high contrast regions, as a strong cue for object boundaries. The white sclera and the dark pupil are optimal cues for edge detection, also because the region is surrounded by low-contrast features. One possibility is therefore, that human glance perception is based on detecting the dark iris and inferring the glance direction from the position of the dark pupil in the white sclera, similar to blob detection algorithms in computer vision [REF].

However, another low-level property of visual processing is that the spatial frequencies, such as skull contours are detected earlier in the process than high frequencies of fine features. While it may seem possible that the receiving mechanism detects and tracks the dark pupil by edge detection, this would not explain the speed and robustness of glance reading seen in humans. Also, blob detection can fail when the object is partly occluded or melts with the environment.

Several high level visual processing modes are commonly known as Gestalts. For example, the Gstalt for symmetry most likely could evolve, because (mirror) symmetry is the most striking feature of almost all extant animals, and therefore is a highly relevant cue. Forming an approximate sphere, the eyeball is strongly symmetric, which is a functional requirement of optical systems, as well as of keeping pressurized fluid in a container.

In most animals, the visual part of the eye is strongly symmetric and two axis, horizontally and vertically. The results of KK suggest that the aspect ratio of the eye region also underlies evolutionary pressure, with longer vertical axes for tree dwellers and longer horizontal axes for ground dwellers. While this is primarily related to the obstruction of the visual field by skull bones (e.g. eye brows), [REF] note that the human eye shape is actually less symmetric than in other Great Apes, when projected on a plane, but highly symmetric, when projected on the skull and observed frontally.

## Bioconvergent eye tracking

If the eye ball sends such strong signals that humans can still read glance directions in highly degraded images, then it must be possible to build an effective eye tracking device that uses the same information. While not guaranteed, it is possible that designing an eye tracking algorithm under these constraints results in a good model, how human glance reading works.

From the characteristics of human glance reading, the following *bioconvergence requirements* can be derived for the algorithm:

1. effective in degraded images and from a distance
2. universal, requiring only minimal calibration
3. be computationally efficient

Most modern eye tracking devices are in so far bio-convergent in that they use cameras and use visual information to predict the glance position. One of the oldest and still very common approaches is to use internal corneal reflections. On early devices the computation was even done with discrete circuits, which made them very fast and efficient. However, corneal reflections are a faint signal which useless over a distance and easily over-ruled by external reflections. This is why the method requires a strong infrared transmitters to create controlled reflections. This is certainly not how humans perceive glance direction. A modern alternative is to use blob detection methods from computer vision. While the pupil is a much larger target than corneal reflections they are computationally more complex and demanding, especially, because of the partial occlusion of the pupil by the eye lids.  
The initial idea for an algorithm that is robust under visible light, computationally efficient and works well with a degraded image over a distance came from the bi-directional symmetry of the visible eye ball, and the idea of an extremely low resolution camera.

The QuadBright method uses the simple fact that the moving pupil produces a change in *spatial brightness distribution*, horizontally and vertically. The image captured by the camera is reduced to a 2x2 matrix in 8 bit greyscale, by splitting it horizontally and vertically in four quadrants (NE, SE, SW and NW, see Fig X) and taking the average brightness (Br). This is used as input for a multiple linear regression model, which predicts the horizontal and vertical position of the eye ball. The following equation shows the model for the horizontal position

$$
h
= \beta\_{h,0}
+ \beta\_{h,\textrm{NE}} \textrm{Br}\_\textrm{NE}
+ \beta\_{h,\textrm{NW}} \textrm{Br}\_\textrm{NW}
+ \beta\_{h,\textrm{SE}} \textrm{Br}\_\textrm{SE}
+ \beta\_{h,\textrm{SW}} \textrm{Br}\_\textrm{SW}\\
v
= \beta\_{v,0}
+ \beta\_{v,\textrm{NE}} \textrm{Br}\_\textrm{NE}
+ \beta\_{v,\textrm{NW}} \textrm{Br}\_\textrm{NW}
+ \beta\_{v,\textrm{SE}} \textrm{Br}\_\textrm{SE}
+ \beta\_{v,\textrm{SW}} \textrm{Br}\_\textrm{SW}
$$

While this In this implementation, the model is trained on Quadbright data on nine calibration points shown initially to the participants. Before every single trial a quick calibration is performed to compensate for minor drifts by adjusting the parameters and .

### YET Zero prototype

Given that QuadBright essentially compresses the input into a four pixel frame, a high resolution camera is not necessary. More important is a small footprint, as it had to be mounted in frontal position. The choice fell on a type of commercially available USB endoscope cameras. These cameras with a diameter of 5.5mm create almost no obstruction when mounted in the visual field and the built-in LED lights provide a stable light source for better accuracy. On the downside these devices deliver a poor resolution of 480x320 and use unspecified electronic components. A simple 3D-printed part was created to be able to glue the camera to a stick, which in turn is connected to a head mount.

The YET Zero application was written in Python, using PyGame for the GUI. The regression model was implemented using the scikit-learn library. First the user is asked to adjust the camera position to be approximately centered. An eye detection algorithm with Haar Cascade is used to initialy detect the eye region, however this is a convenience feature in this implementation is not related to the actual eye tracking.

This system already worked well, but it was susceptible to changes in computer screen brightness effects introduce by different stimuli. While the best solution is to use a frontal brightness sensor and extend the model accordingly, as a quick fix the quick calibration was augmented with a highly degraded preview of the next stimulus, just enough to keep the brightness level stable.

## Results

## Study 1: Accuracy of the QuadBright algorithm

The first study was conducted to measure the accuracy of the Quadbright algorithm, with the expectation that it matchges human performance. In the experiment participants were asked to follow a target, while the QuadBright algorithm tracked their eye positions. Accuracy was measured by calculating the angular accuracy of the predicted positions.

The first study was conducted to evaluate the accuracy of the QuadBright algorithm. The algorithm was tested on a set of 36 participants, who were asked to find and focus an orange target on a white background, moving in a random pattern between 24 positions.

The experiment was carried out with Yet Zero. The headmount was constructed from an old headphone and a kitchen paper roll was used as a chin rest. Participants were seated in 45cm viewing distance using a 16:9 screen with a 41cm diagonal. The experiment consisted of 24 trials, each lasting 3 seconds. The target moved in a random pattern between 24 positions on the screen. The participants were asked to follow the target with their eyes. The QuadBright algorithm tracked their eye movements and the accuracy of the algorithm was measured by calculating the angular standard error of predicted eye movements of the participants.

The original experiment tested various lighting conditions (LED, IR, no light), as well as the before-mentioned bias due to changing screen brightness levels ((FB24)). The latter was confirmed to such an extent that the results were discarded. The three levels of lighting conditions differed so little that for the analysis here data was pooled.

FB24\_1 <-   
 FB24 |>   
 group\_by(Part, target\_x, target\_y, trial) |>   
 summarize(accuracy = mean(visual\_angle\_accuracy)) |>   
 ungroup()

`summarise()` has grouped output by 'Part', 'target\_x', 'target\_y'. You can  
override using the `.groups` argument.

FB24\_1 |>   
 group\_by(target\_x, target\_y) |>  
 summarize(mean\_error = mean(accuracy), sd\_error = sd(accuracy), median = median(accuracy))

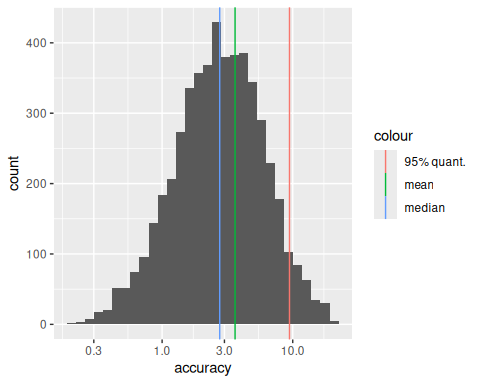
`summarise()` has grouped output by 'target\_x'. You can override using the  
`.groups` argument.

# A tibble: 24 × 5  
# Groups: target\_x [6]  
 target\_x target\_y mean\_error sd\_error median  
 <dbl> <dbl> <dbl> <dbl> <dbl>  
 1 160 135 6.22 4.82 5.36  
 2 160 405 5.91 4.20 5.16  
 3 160 675 5.75 3.93 4.98  
 4 160 945 6.24 4.87 5.21  
 5 480 135 4.22 2.98 3.55  
 6 480 405 3.77 3.35 2.87  
 7 480 675 3.34 3.10 2.57  
 8 480 945 3.71 3.43 2.89  
 9 800 135 3.95 4.13 2.92  
10 800 405 2.52 2.21 1.86  
# ℹ 14 more rows

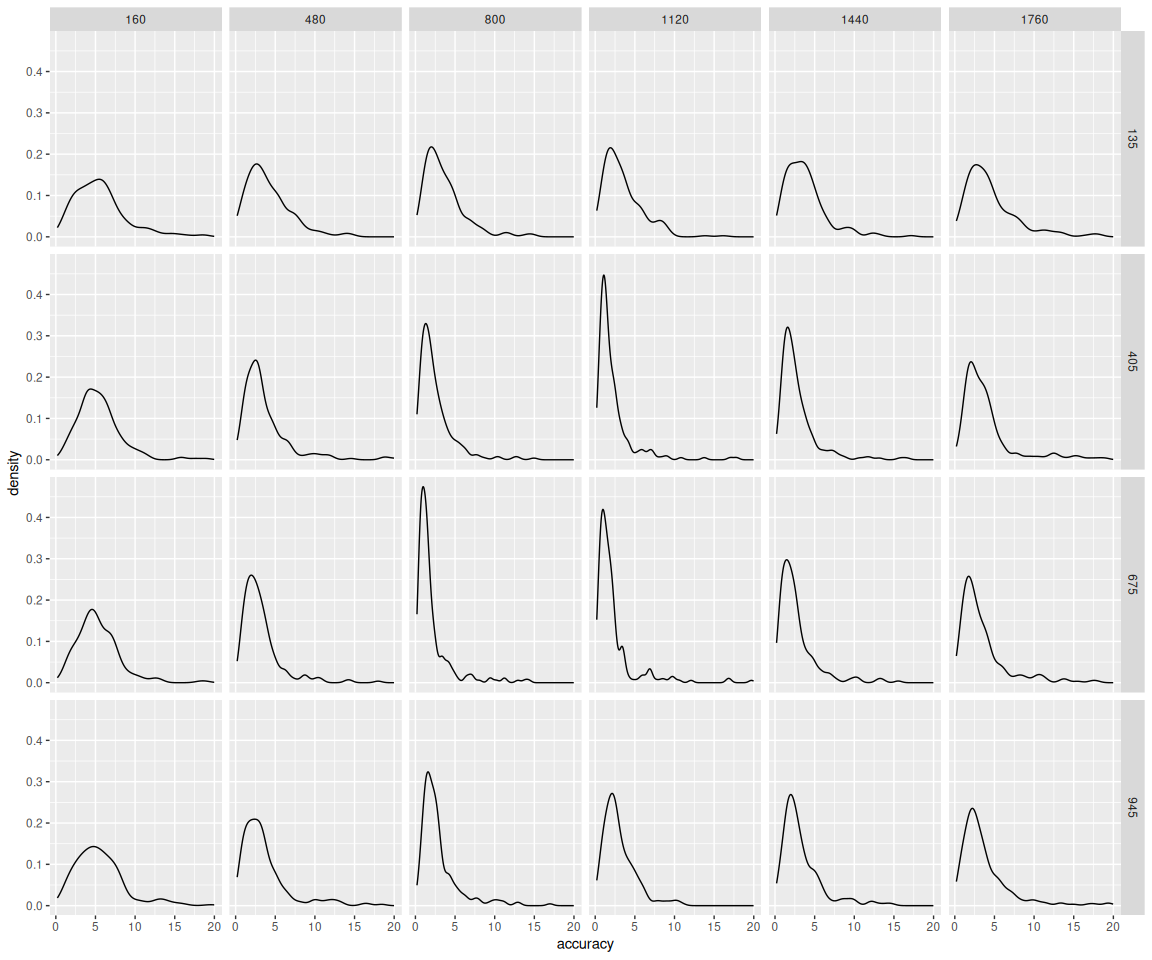
Figure ((XZ)) shows the distribution of angular errors combined (on scale). The median accuracy is 2.8 degree, with a mean of 3.2 degree. 95 percent of all errors are below 8.5 degree. The accuracy is best in the center region with a median of 1.3 - 1.4 degree accuracy. The error increases towards the edges of the screen, but asymmetrically. To the mount side the median error is 2.6 - 3.7 degree, but 5 - 5.5 degrees to the inner side of the eye. In all positions, errors larger than 10 degree were extremely rare. We can conclude that the QuadBright algorithm is able to estimate eye ball positions reasonably accurate and robust under different light conditions. The few catastrophes mostly occurred on three (out of 36) participants, for which the calibration didn’t seem to work reliably.

T\_0 <-   
 FB24\_1 |>   
 summarize(mean(accuracy), median(accuracy), quantile(accuracy, .95))  
   
  
FB24\_1 |>   
 filter(accuracy < 20) |>   
 ggplot(aes(x = accuracy)) +  
# facet\_grid(target\_y ~ target\_x) +  
 scale\_x\_log10() +  
 geom\_histogram() +  
 geom\_vline(aes(xintercept = median(accuracy), col = "median")) +  
 geom\_vline(aes(xintercept = mean(accuracy), col = "mean")) +  
 geom\_vline(aes(xintercept = quantile(accuracy,.95), col = "95% quant."))

`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



FB24\_1 |>   
 filter(accuracy < 20) |>   
 ggplot(aes(x = accuracy)) +  
 facet\_grid(target\_y ~ target\_x) +  
 geom\_density()



It can be concluded that the QuadBright algorithm is able to estimate eye ball positions reasonably accurate and robust under different light conditions. The few catastrophes mostly occurred on three (out of 36) participants, for which the calibration didn’t seem to work reliably. This is a common problem to all eye tracking devices.

## Study 2: Human glance perception with QuadBright information

The first study established that a simple machine learning algorithm can estimate eye positions with a reasonable accuracy from highly degraded images. The Quadbright algorithm is therefore a viable candidate model for human glance perception. The second study was conducted to test this possibility by assessing human glance reading performance on images degraded by the Quadbright method.

### Experimental design

In this experiment, participants see a frontal face glancing at the hours positions on a clock face (Fig X a) and are asked to read the correct hour position. For the experimental condition, eye regions were reduced to four pixel according to the QuadBright method (Fig X b). To test the robustness, exposure times were varied in five steps (1000, 600, 400, 140 and 70 milliseconds).

### Sample

### Data analysis

The outcome variable is the number of deviations between true and reported hour positions, resulting in a count variable. A two-factorial linear term with conditional effects of exposure time and QuadBright degradation was used for population-level (fixed effects), as well as on Stimulus and Participant level (random effects). To account for over-dispersion, a negative-binomial outcome distribution with logarithm link function was tried first, but the extremely large reciprocal dispersion parameter suggested that the data is not over-dispersed and the final model was therefore fitted using a the Poisson family.

### Results

library(tidyverse)  
library(rstanarm)  
library(tidyverse)  
  
options(mc.cores = 8)  
  
JGJH <-   
 readxl::read\_excel("JGJH/JGJH.xlsx") |>   
 mutate(Obs = row\_number(),  
 Part = as.factor(participant\_number),  
 Exposure = fct\_rev(as.factor(presentation\_time\_ms)),  
 Stim = as.factor(stimulus)) |>  
 select(Obs, Part, Stim, Condition = condition, Exposure, deviation)  
  
M\_1 <- stan\_glmer(deviation ~ Condition \* Exposure + (1 | Part), family = poisson, data = JGJH)  
M\_2 <- stan\_glmer(deviation ~ Condition \* Exposure + (Condition | Part), family = poisson, data = JGJH)  
  
M\_4 <- stan\_glmer(deviation ~ Condition \* Exposure +   
 (Condition \* Exposure | Part) +  
 (Condition \* Exposure | Stim),   
 family = poisson(), data = JGJH,  
 chains = 8)  
  
save(JGJH, M\_1, M\_2, M\_4, file = "Models.Rda")

library(bayr)

Attache Paket: 'bayr'

Das folgende Objekt ist maskiert 'package:tidyr':  
  
 expand\_grid

load("Models.Rda")

Registered S3 methods overwritten by 'rstanarm':  
 method from  
 coef.stanreg bayr  
 predict.stanreg bayr

P\_4 <-   
 posterior(M\_4, thin = 5) |>   
 mutate(value = if\_else(type %in% c("fixef", "ranef"), exp(value), value),  
 value = if\_else(fixef == "Intercept", value \* 30, value))

The model uses treatment effects coding with Intercept being the baseline (original image, 1000ms). Table X shows population-level coefficients transformed to degrees of deviation, with multiplicative treatment effects. With original images and 1000ms exposure, average deviation is around 13 degree, albeit with some uncertainty. There is quite some variation in participants, and even more so in stimuli.

It seems that individual differences exist in the ability to read glances, whereas the variation in stimuli

For 600 - 400ms exposures performance , all coefficients are very close to 1, which means that performance remains stable. However, levels of certainty are moderate, such that a slight decrease in performance cannot be ruled out. With 140ms, performance drops slightly in both conditions, with around 20% stronger deviations. With 70ms, performance remains stable in the original condition, but deviations increase by almost 30% with degraded images.

With degraded images, deviations increase by around 70%.

fixef\_ml(P\_4)

Population-level coefficients with random effects standard deviations

| fixef | center | lower | upper | SD\_Part | SD\_Stim |
| --- | --- | --- | --- | --- | --- |
| Intercept | 13.2653663 | 10.3724979 | 16.788733 | 5.2725142 | 11.4535789 |
| Conditionexperimental | 1.6995089 | 1.3745957 | 2.115061 | 0.3130282 | 0.3002775 |
| Exposure600 | 1.0561627 | 0.9372393 | 1.191474 | 0.0689191 | 0.0806759 |
| Exposure400 | 1.0684788 | 0.9532992 | 1.198129 | 0.0506691 | 0.0718387 |
| Exposure140 | 1.1845632 | 1.0536886 | 1.337245 | 0.1190388 | 0.0905790 |
| Exposure70 | 1.1900177 | 1.0526121 | 1.355623 | 0.1267368 | 0.1019595 |
| Conditionexperimental:Exposure600 | 1.0343328 | 0.8968198 | 1.210703 | 0.0858580 | 0.0935776 |
| Conditionexperimental:Exposure400 | 0.9990469 | 0.8729574 | 1.163019 | 0.0738007 | 0.0746266 |
| Conditionexperimental:Exposure140 | 0.9845297 | 0.8528147 | 1.146573 | 0.1072058 | 0.0843402 |
| Conditionexperimental:Exposure70 | 1.2886261 | 1.0859864 | 1.521978 | 0.1869846 | 0.1618428 |

## Discussion

The Cooperative Eye Theory states that the eyeball has evolved to be a sender of glance direction signals. If that is true, it must be easy to construct a machine algorithm reading glance directions, as the human visual system is able to do so. We showed that a remarkably simple algorithm can estimate eye positions with good accuracy, using four-pixel images and a 10-parameter model. This model can be solved by the method of least-squares, where all computations during training and prediction are analytically closed, requiring no costly numerical approximations, recursions, or iterative procedures. Memory consumption is minimal, too, requiring only 14 Byte for making a prediction.

Additionally, it was required that the algorithm must function under natural conditions, such as reading glances from a distance. We did not test this directly, but the mere fact that glance directions can be estimated from a 2x2 pixel matrix suggests that the algorithm could definitely work. The Quadbright method has even more virtues that enhance its utility and may therefore have (had) additional reproductive advantage. The little information it needs also works in the low-resolution is not just present in the fovea region, but also in the peripheral vision and even in night vision. If the group-in-action scenario proposed by KK is the origin of glance reading, then it is more useful if the receiver doesn’t have to look directly at the sender, but can read glance directions from the periphery. This may also explain the common feeling of “someones eyes on me”. For hunters it is also very useful to be able to read glance directions in low light conditions, such as at dusk or dawn.

Altogether, the eye ball as a sender and the Quadbright algorith on the receiving end seem like a very good fit. However, to truly work under real circumstances, several other components are needed, for the eye tracker, but also in the human mind. In real situations the image is not stable, but constantly changing. Another feat of the human visual system is how it can track objects, i.e. the eye region, with ease. The oculomotor system even has a special mode for this, called smooth pursuit [XY]. Object tracking is a feature in many technical systems, such as autofocus for cameras, or assisted driving systems.

The eye tracker is relatively sensitive to changing lighting conditions, which was mitigated by a quick calibration procedure before each trial. However, the human visual system is extremely adaptive to lighting conditions [XY], even unmatched by commonly available sensor technology. This also regards spatial shifts in brightness, as shadow-forming even is an important cue for depth perception. A simple way to emulate this ability would be to add another 2x2 pixel camera to the front of the system to capture the incoming light distribution. This would add another four parameters to the model.

Human glance reading also takes the senders head position into account. This has been shown to be the major mechanism for shared attention detection in chimpanzees, so it either a very old cognitive function, or it is convergent. Head position is commonly based on algorithms that first detect a configuration of prominent facial features and then estimate the head position from the relative positions of these features. To emulate the human visual system, another camera could be used, facing the participant from a fixed position. The Quadbright algorithm could be extended to include a head position estimate in much the same way, but technically is probably easier, and more efficient, to use the signal of a gyroscope sensor attached to the head mount.

By its existence, the Quadbright method proves that the human eye ball has evolved to send a robust and computationally efficient glance direction signals. Overall, glance reading accuracy in this experiment was surprisingly low with original images (around 14 degree on average), where others have reported much higher accuracy [XY]. A possible cause is that the experiment used quite extreme target positions. It is known that glance reading performance deteriorates to the periphery [XY], and also the eye tracker shows lower accuracy in extreme positions.

At the same time, signal efficiency of the eyeball is very evident by the stable performance down to 400ms exposure. Only at 140ms performance accuracy drops slightly, but then remains stable even down to 70ms. This supports the idea that the eye ball has evolved for computational efficiency.

But, is the Quadbright algorithm bio-convergent on the deeper level of cognitive processing? Is the visual system solving a series of simple linear equations? If the algorithm is a good model of human glance perception, then humans should be able to read glance directions from the same information available to the eye tracker, with similar accuracy.

What we observed is that humans can definitely read glance directions from the degraded images, albeit with a moderate loss in performance compared to the baseline condition. This offset is stable down to 140ms exposure. Only at 70ms exposure, performance drops by around 20% compared the original image. While the overall similarity is striking, the anomaly at 70ms can possibily explained by the fact that we refrained from visual masking to over-write iconic memory. The iconic memory acts as a low-level visual buffer, which prolongs the available processing time for higher-up recognition mechanisms. Image decay in iconic memory is known to be slower with higher contrasts images [REF], which in this case is the original image.

Given the reduced performance, we cannot conclude that the Quadbright algorithm is a perfect model of human glance perception. However, since participants were able to read glance directions to some extent, we can rule out the hypothesis that human glance perception is based on blob detection, as this information is simply not available in the degraded condition.

We argued that from en evolutionary perspective, a simple algorithm is more likely to develop than a complex one, especially given the relatively short period in which it could develop. The Quadbright approach takes this idea to an extreme, in that a 2x2 pixel matrix is the bare minimum to estimate vertical and horizontal glance directions. It neglects the observation that the human glance directions are dominantly horizontal, which is expressed in the characteristic shape of human eyes. Future studies should therefore explore a slightly enhanced algorithm using a 3x2 pixel matrix, which would allow to estimate the horizontal glance direction more accurately, with only a small increase in computational complexity.

Finally, it appears that the eye tracker had a better overall accuracy as human participants. This may also suggest that the cognitive mechanism is not fully optimized, which supports our theoretical conclusion that it must be relatively young in evolutionary time scales.

## Future research

The results suggest that brightness distribution models are a good starting point in understanding cognitive glance processing. However, we only showed that performance is in the same ball park. More refined studies should analyze more closely the systematic biases that occur with brightness distribution algorithms and test whether the same bias can be provoked in human participants.

As an example, the large area reflections from the computer screen induced strong offsets in glance directions predicted by the eye tracker, and it can be tested whether similar biases occur in cognitive glance reading. At time of evolution, shiny large surfaces, such as computer screens or sheets of paper were present to a much lesser extent, so we can expect that the cognitive system has not evolved to deal with these situations.

In more general, human glance reading must include other sources of information, such as head position, and it is to date unknown how the cognitive system integrates these sources of information.

## Applications

While the primary aim of this work is to test the CEH and postulate the reciprocal evolution of a receiver mechanism, there are opportunities to exploit the results for practical applications. First, we could demonstrate that building an effective eye tracking device is remarkably easy. While the pilot implementation does not reach the same level as commercial eye trackers, the differences are not as large as one would expect, at a price point reduced by two magnitudes. The YET Zero prototype is already used routinely in bachelor-level classes, and further improvements, such as using infrared transmitters could improve the performance further.

Another field of application is the design of artificial faces. In social robotics, robot faces are often designed to be expressive, and the role of glance reading in human-robot interaction is well established. However, designing more human-like eye regions for this purpose may result in an adverse effect known as the Uncanny Valley. Artificial faces that are too human-like, but not quite right, are perceived as creepy. The results of this study suggest that a four-pixel eye region is sufficient to convey glance direction, which could facilitate the quality of interactiuon, without falling into the Uncanny Valley.

## Limitations