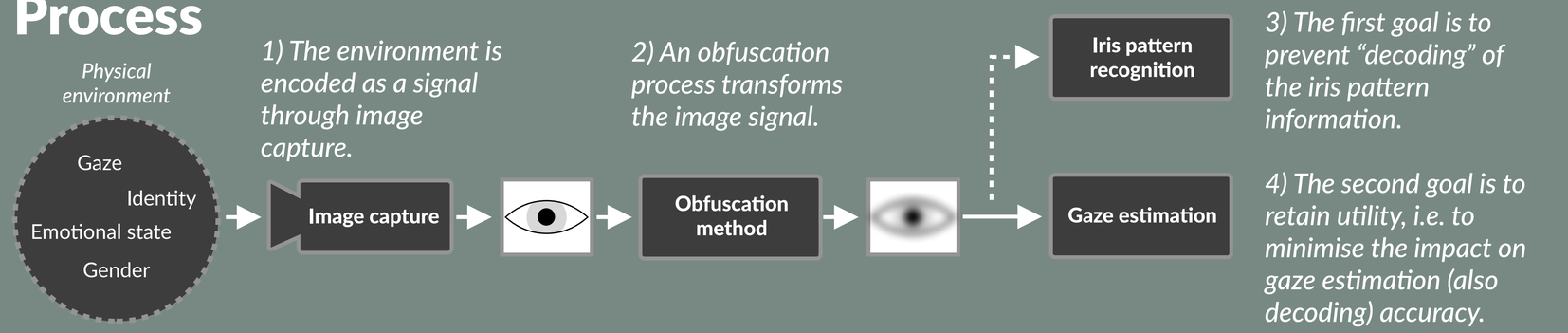
Eye images can be modified to remove **iris pattern information** while remaining usable for **gaze estimation**

Terms

Obfuscation: The process of making an information source unintelligible.

Utility: The quality of a data sample with respect to a specific purpose.

Process



Analysis of iris obfuscation: Generalising eye information processes for privacy studies in eye tracking.

Anton Mølbjerg Eskildsen, Dan Witzner Hansen

In this study we:

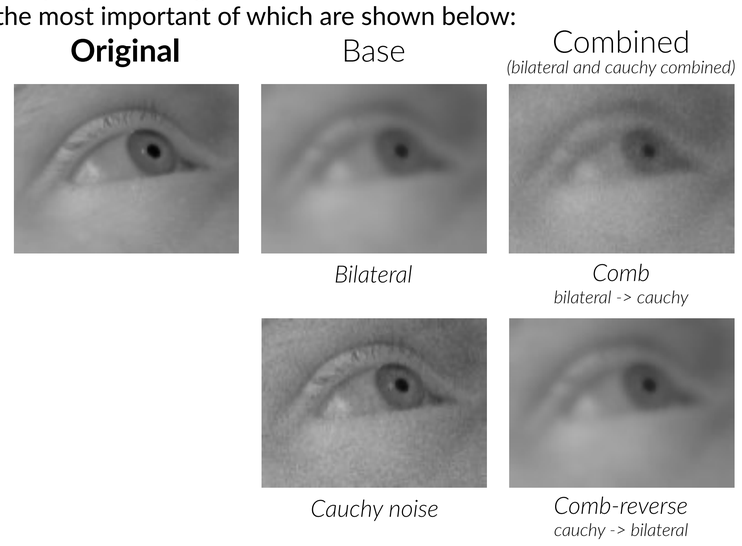
- Present methods that effectively prevent iris recognition while retaining utility for gaze estimation.
- Propose a framework that defines obfuscation and utility as general concepts in eye-tracking.

INTRODUCTION

Iris patterns are stable biometrics that can be used for personal identification. Image-based eye-trackers capture high-quality eye images that may be used for iris recognition. This is undesirable for legal and ethical reasons since this sensitive data is always present. We propose obfuscation methods that remove the iris pattern information while retaining high utility for gaze estimation.

METHOD

We propose a large number of candidate methods for obfuscation, the most important of which are shown below:

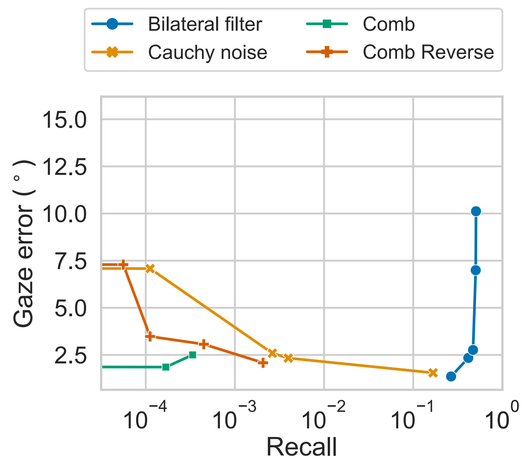


Parameter optimisation:

- Optimal parameters are found using grid search and selecting Pareto-optimal points (We use the CASIA-V4 dataset for all experiments [4]).
- Precise security was measured as the probability of an attacker finding a correct match given favourable conditions*. Specifically, we find the maximum recall for a precision of 1.

RESULTS

Worst-case attacker results. Lower recall is better (worse for the attacker) and lower gaze error is better. Lines trailing off to the left indicate recalls of 0**:



CONCLUSIONS

- Obfuscation using the proposed methods is generally effective at preventing iris recognition using an iris-recognition system of reasonable precision derived from [1]. This is a significant improvement over previous studies [2, 3].
- Probabilistic attacks are still possible, as shown in the graph above.
- The framework introduced presenting obfuscation and utility in terms of notions from information theory is shown to be effective for understanding, evaluating and proposing similar systems that aim to remove information in this manner.
- Future work is aimed at deep learning based models and working towards more definitive measures for determining risks.

REFERENCES

[1]: Daugman, J. "How Iris Recognition Works". *IEEE Transactions on Circuits and Systems for Video Technology* 14, nr. 1 (januar 2004): 21–30. https://doi.org/10.1109/TCSVT.2003.818350.

[2]: John, Brendan, Sanjeev Koppal, og Eakta Jain. "*EyeVEIL*: degrading iris authentication in eye tracking headsets". I *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*, 1–5. ETRA '19. New York, NY, USA: Association for Computing Machinery, 2019. https://doi.org/10.1145/3314111.3319816.

[3]: John, Brendan, Ao Liu, Lirong Xia, Sanjeev Koppal, og Eakta Jain. "Let It Snow: Adding pixel noise to protect the user’s identity". I *ACM Symposium on Eye Tracking Research and Applications*, 1–3. ETRA '20 Adjunct. New York, NY, USA: Association for Computing Machinery, 2020. https://doi.org/10.1145/3379157.3390512.

[4]: CASIA-Iris-Internal-V4. http://biometrics.idealtest.org/

OPTIMISING FOR OBFUSCATION AND UTILITY

$$\arg\min_{O} \quad \mathcal{E}_a^O$$

$$\text{subject to} \quad \mathcal{E}_b^O \leq l$$

Utility function:

$$\mathcal{E}_{util}^{\mathcal{O}} = \sum_{i=1}^{n} ||f_{gaze}(\mathcal{O}(I_i)) - p_i||_2$$

Obfuscation measure:

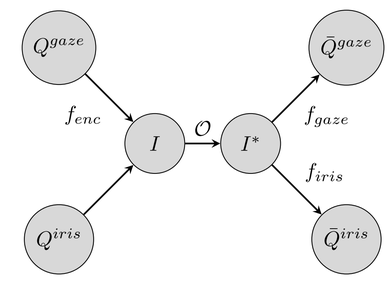
$$\mathcal{E}_{obf}^{\mathcal{O}} = -\mathbb{E}[h(f_{iris}(I), f_{iris}(\mathcal{O}(I)))]$$

Alternative formulation:

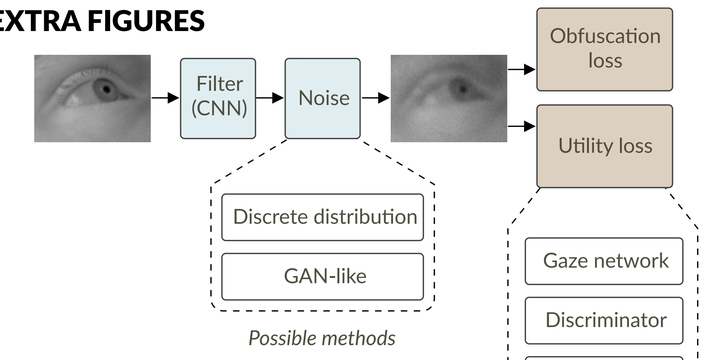
$$\mathcal{E}_{obf}^{\mathcal{O}} = \mathcal{J}(\bar{Q}; Q) - H(\bar{Q} \,|\, Q)$$

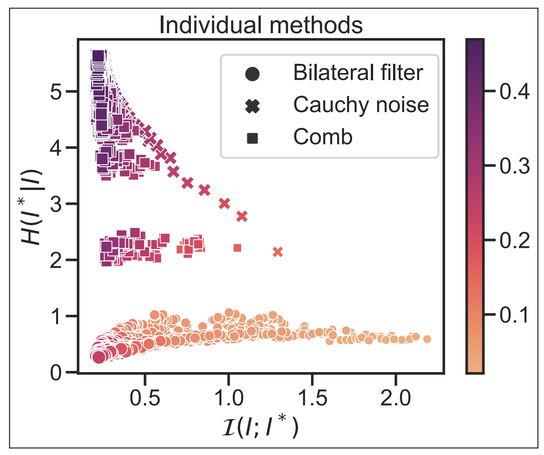
Image entropy

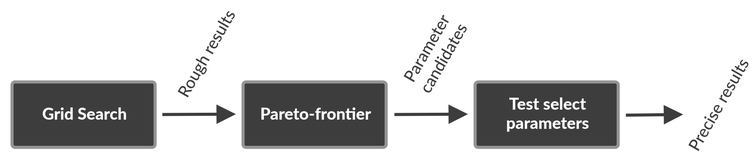
$$P_{\nabla I}(a, b) = \frac{1}{N} \sum_{\mathbf{x} \in \mathcal{X}} \delta_{a,\partial_x(\mathbf{x})} \delta_{b,\partial_y(\mathbf{x})}$$



EXTRA FIGURES







*Attacker has access to a labelled dataset of obfuscated images (image + identity) and can therefore optimise their iris-recognition model to the specific obfuscation method used. More details in the paper.

**The CASIA-V4 dataset only has 17864 possible positive matches when using all image combinations. The recall is therefore limited to a precision of $10^{-4.252}$ which is exactly the minimum value of the plot. In other words, the 0 values are not expected to actually be 0, they are simply below the precision limit. The limit also suggests that results close to it may be imprecise to sample variance.

IT UNIVERSITY OF CPH

Take a picture to **download** the **full paper**