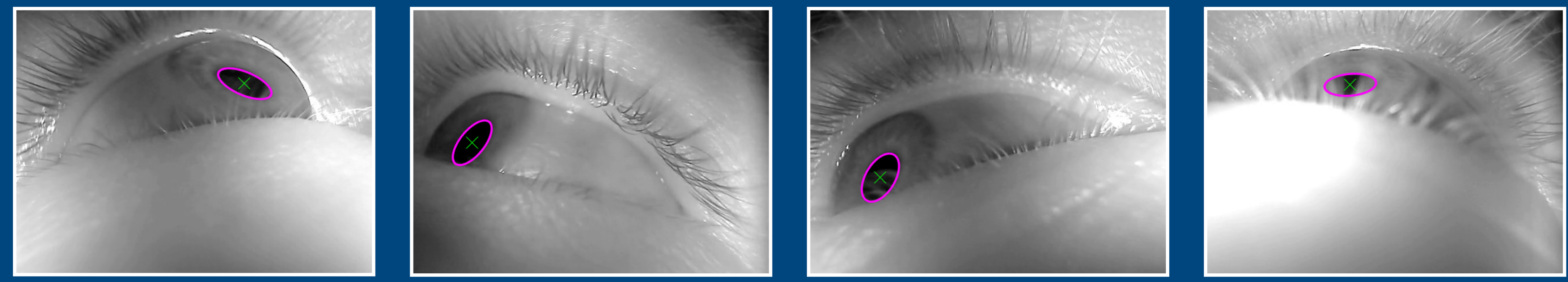
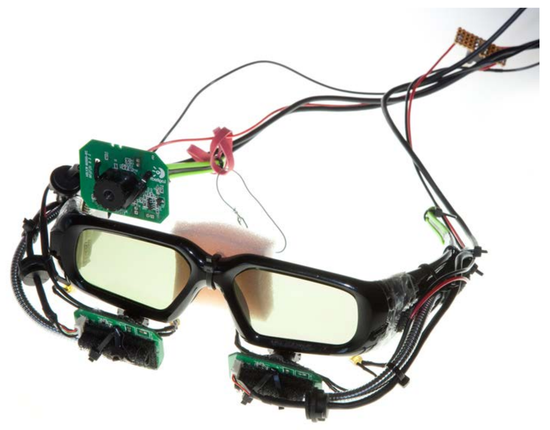
Robust, Real-time Pupil Tracking

Lech Świrski, Andreas Bulling & Neil Dodgson



Motivation

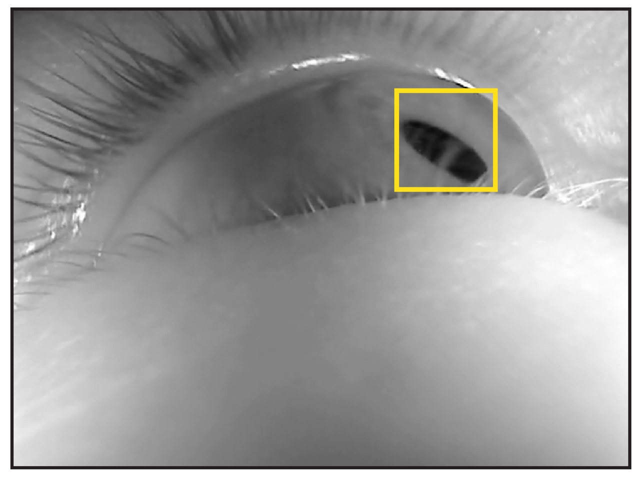
We wanted to mount an eye-tracking camera very close to a user's eye, underneath a pair of glasses. Existing pupil tracking approaches struggled, due to eyelash occlusions and highly elliptical pupils. Our pupil tracking is robust to such issues, and does not rely on hardware invariants (such as fixed lighting), making it suitable for low-cost webcam-based eye trackers.



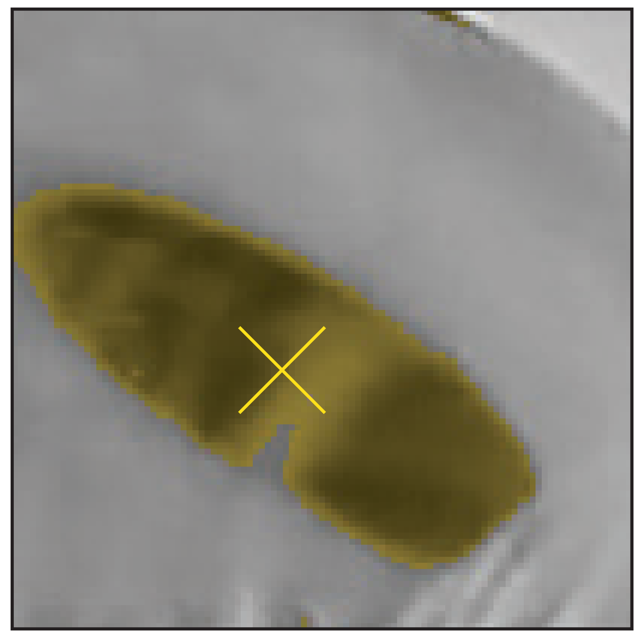
Our eye-tracking cameras, mounted underneath glasses

Overview

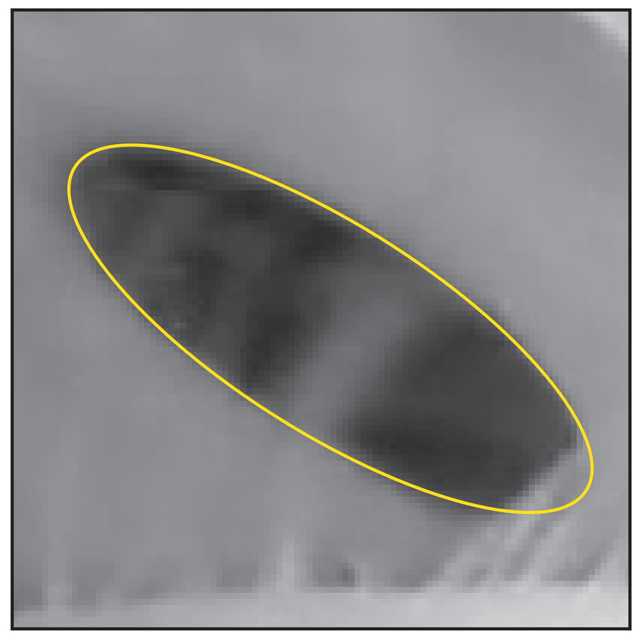
To detect pupils both accurately and in real time, our algorithm cheaply approximates the pupil location in one stage, then uses more expensive algorithms on a smaller region to accurately refine this approximation.



Initial pupil region estimation using a fast, simple feature detection.



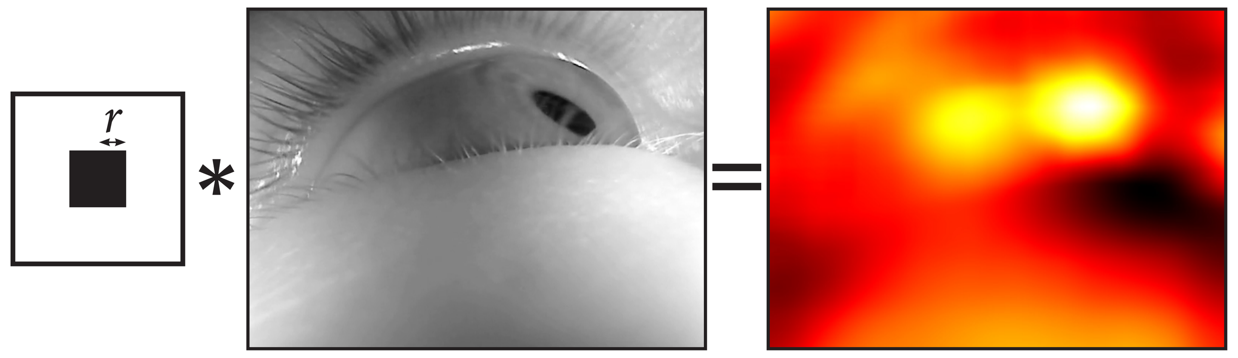
Adaptive threshold-based pupil segmentation to refine the pupil centre approximation



Robust ellipse fitting using a novel, pupil-specific RANSAC algorithm.

Initial region estimation

The pupil region can, very roughly, be described as a "dark blob surrounded by a light grey blob". We detect regions fitting this description using a Haar-like centre surround feature detector. The best such region is selected as the approximate pupil region. Using integral images, this operation is fast and cheap.

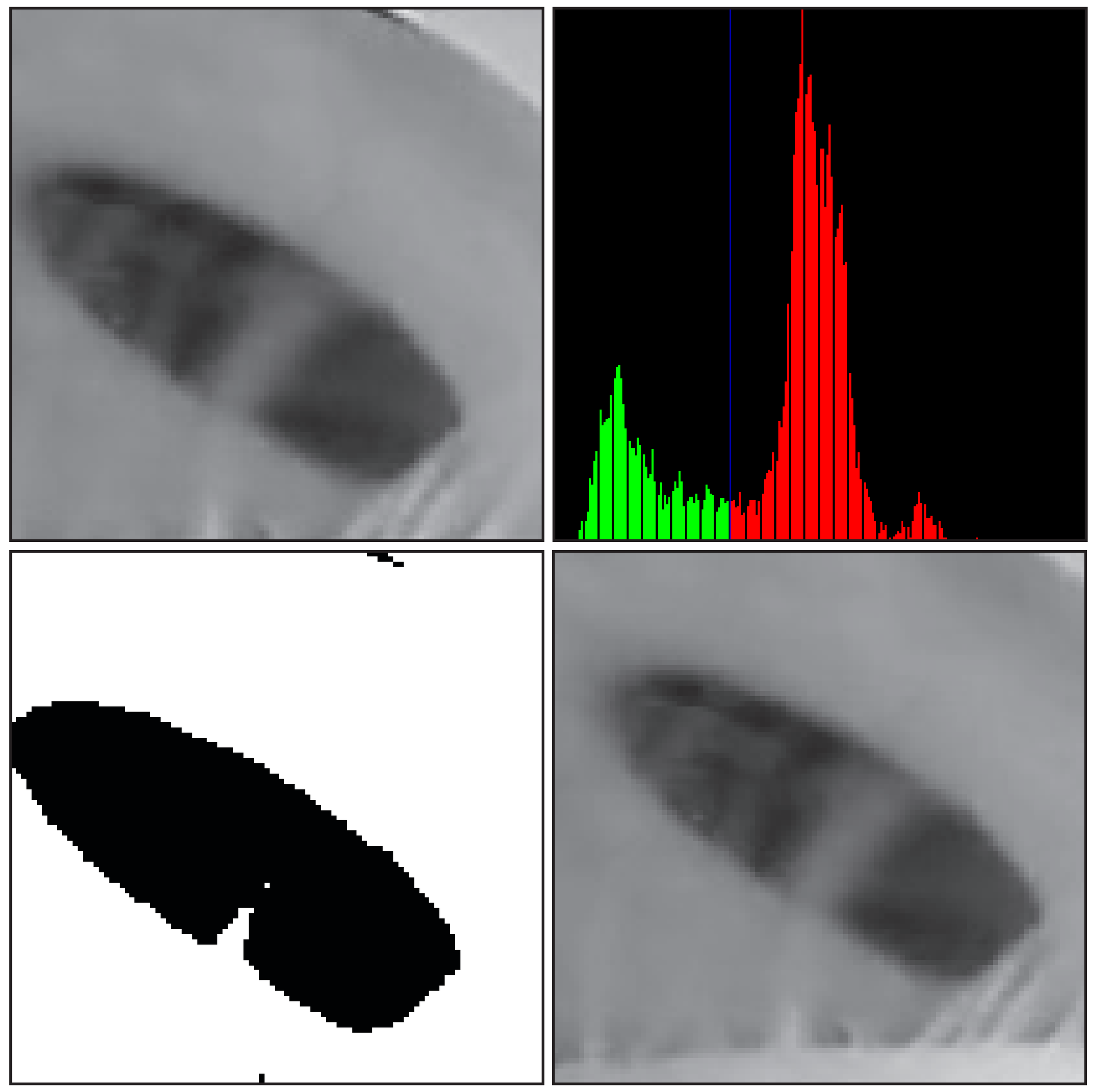


By convolving with a Haar-like centre surround feature, of given "radius" r, we produce a heat map of likely regions. The location of the highest value in the heat maps across all radiuses is assumed to be the centre of the pupil region.

Approximate pupil segmentation

Thresholding is commonly used as a pupil detection technique. Setting the threshold is crucial to the quality of detection, however is often done manually, and is affected by changes in illumination and camera settings.

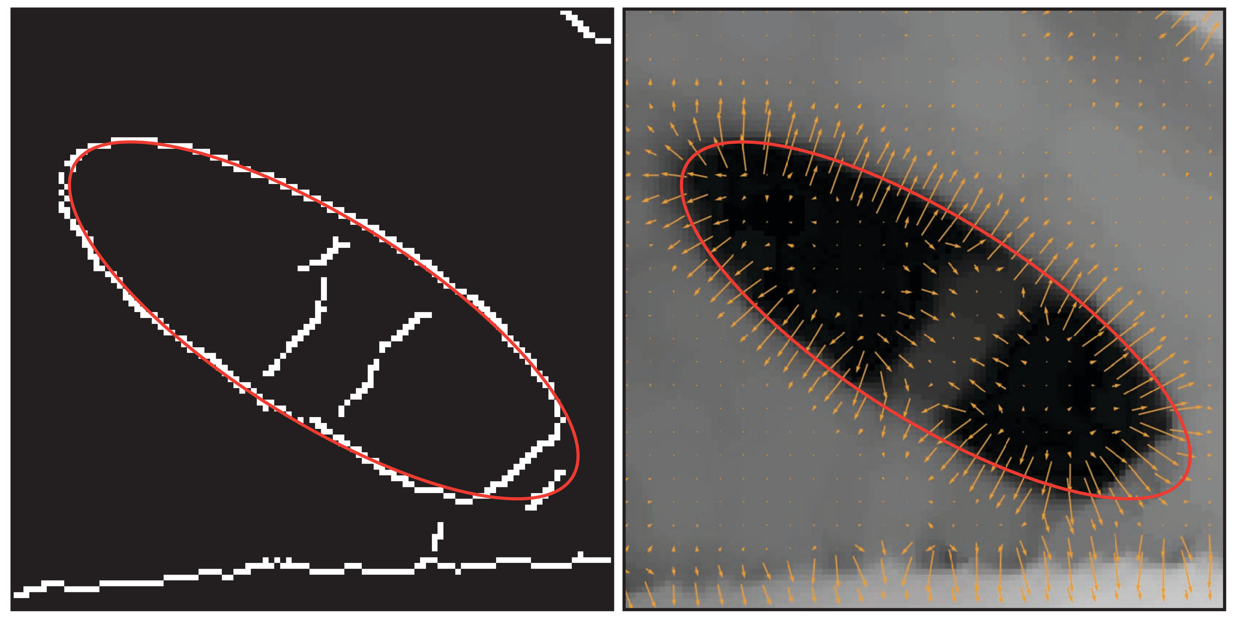
We use k-means clustering on the histogram of the approximate pupil region to detect dark and light clusters, separated by some threshold luminance value. The largest dark component in the thresholded image is assumed to be the pupil, and the pupil region is re-centred on this dark component



The pupil region is thresholded by histogram segmentation. The largest dark pixel component is assumed to be the pupil, and the pupil region is re-centred

Ellipse fitting

To accurately find the pupil centre, we fit an ellipse to the pupil image. The centre of the ellipse is assumed to be the centre of the pupil. We use Canny edge detection to find likely pupil edge pixels. We can use an expensive edge detector as we are only evaluating it in the small pupil region. We then use RANSAC to fit an ellipse to the edge points.

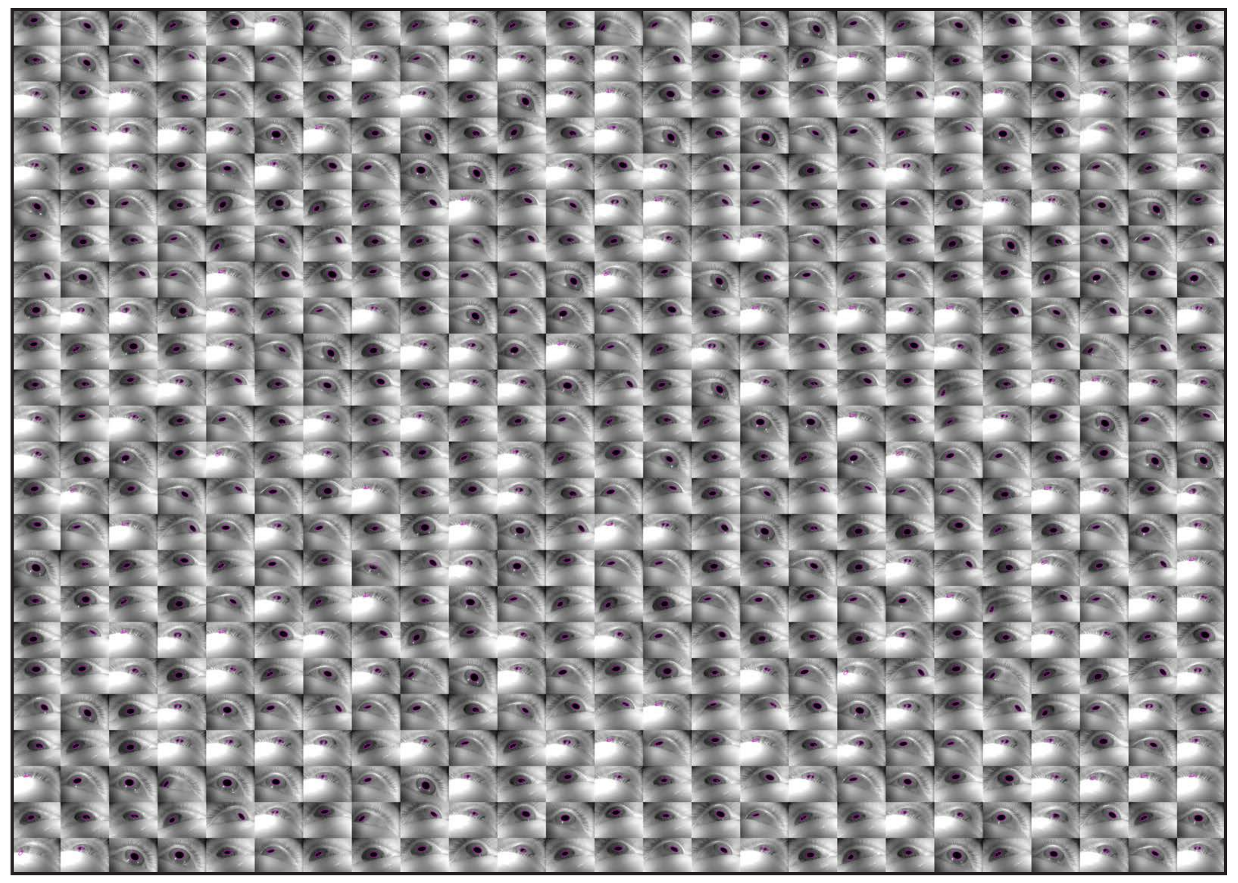


We evaluate the quality of an ellipse fit using both the Canny edges and the gradient field of the image. The best fitting ellipse is assumed to be the pupil.

RANSAC fits an ellipse to a set of points by iteratively fitting ellipses to random subsets of the set, and selecting the best fit. It allows us to be robust to outliers, such as edges of eyelashes and eyelids. We use a novel, pupil-specific ellipse fit metric, which tries to find an ellipse that matches the edge image points and is orthogonal to the gradients of the image.

Evaluation

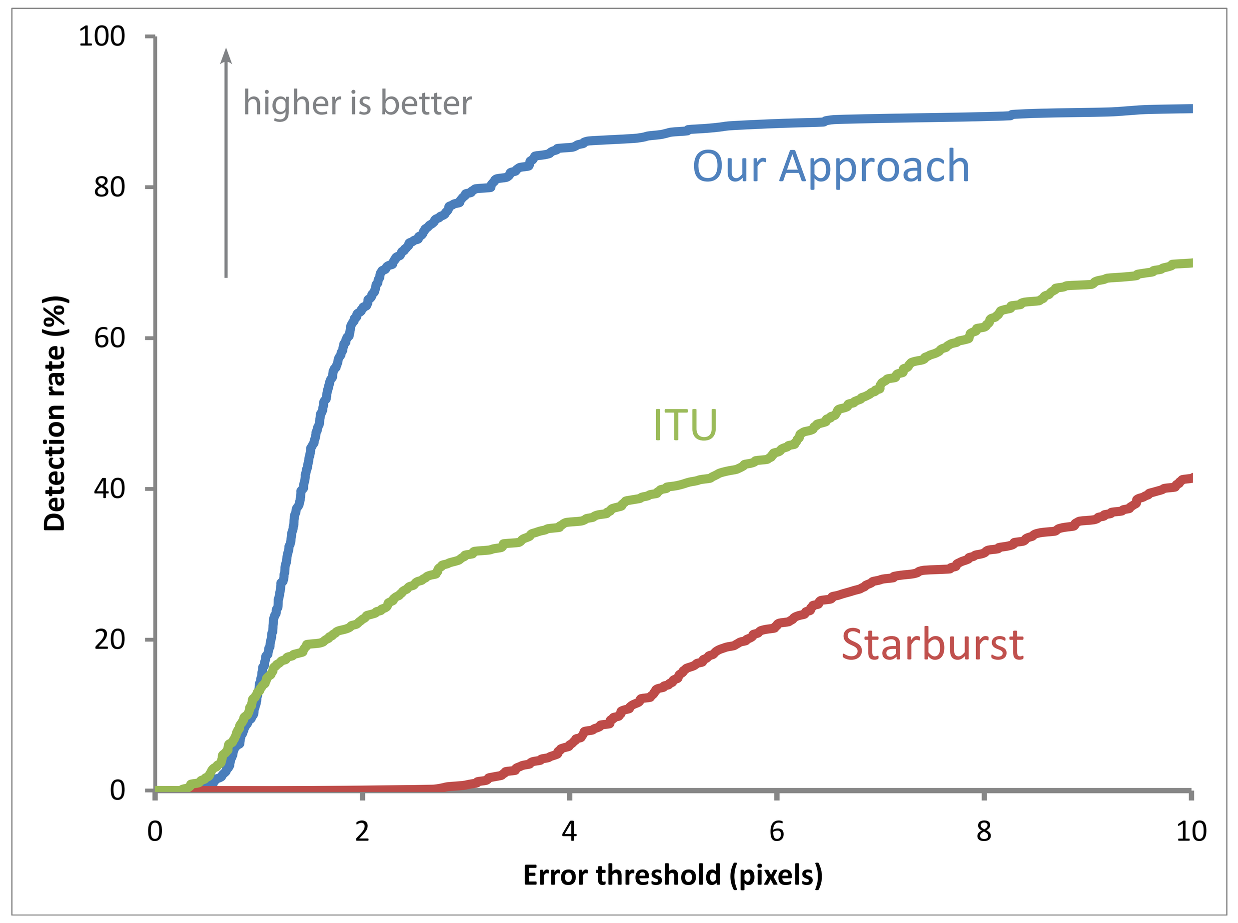
We evaluate our pupil detection algorithm by running it on a dataset of eye images, and comparing the pupil ellipses to a ground-truth. We created this dataset by recording two people's left and right eyes, and manually labelled the ellipses on a 600 frame subset.



We manually labelled the pupil ellipses in 600 eye images to use as the ground truth for the evaluation.

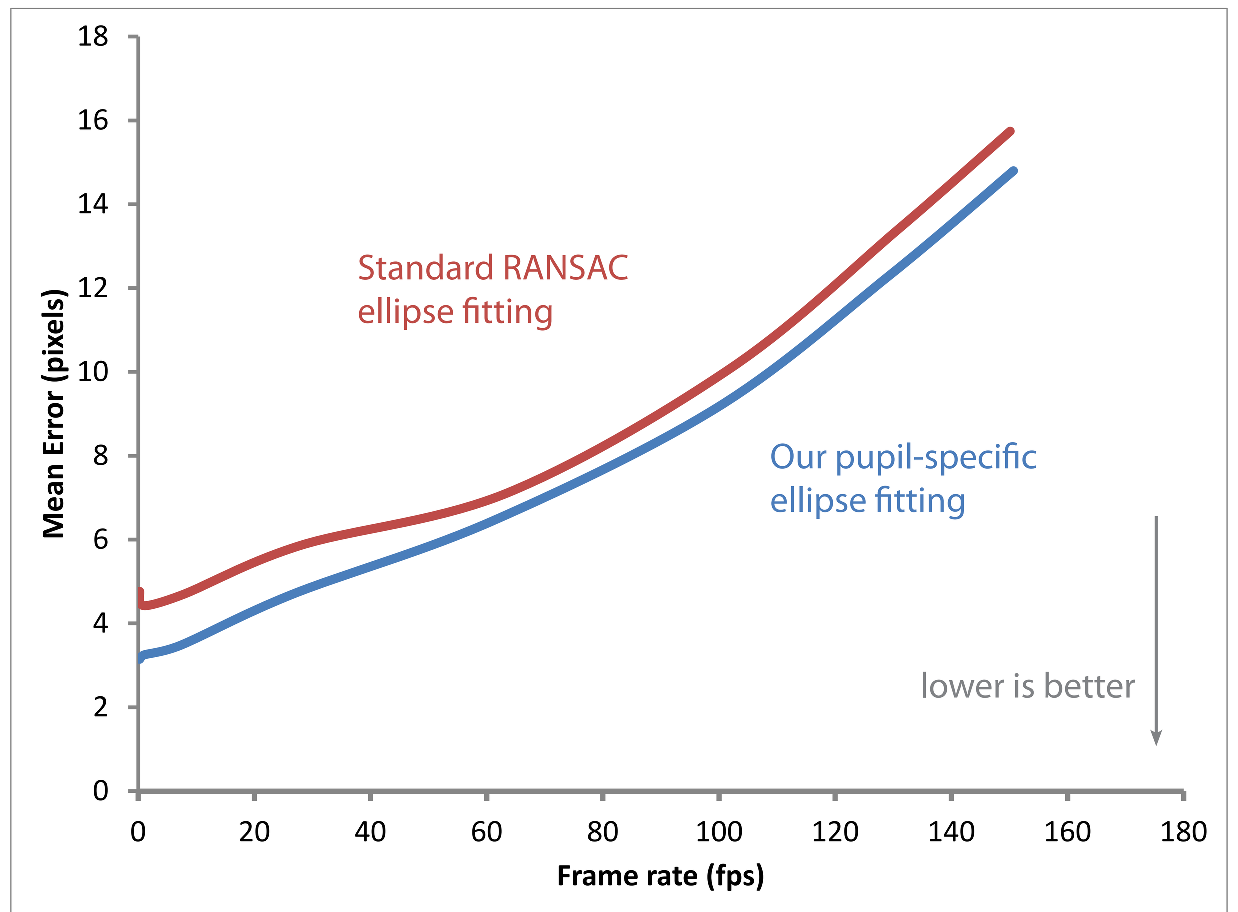
Results

We compared our pupil detection algorithm to two existing real-time pupil trackers: Starburst and the ITU Gaze Tracker. We measured the detection rate of each technique by finding the percentage of pupil ellipses correctly identified within some error threshold.



The pupil detection rate of our approach compared to Starburst and the ITU Gaze Tracker. We achieve a detection rate of over 80% within a 4px error.

We also evaluated the accuracy of our method for various framerates, by varying the maximum number of iterations in the RANSAC ellipse fitting stage. We found that, for real-time speeds (30 fps), our pupil tracking algorithm has a mean error of less than 5px. We also found that our pupil-specific ellipse fitting universally improves on standard RANSAC ellipse fitting by approximately one pixel.



The mean error of our pupil tracker running at various framerates, with and without our novel pupil-specific ellipse fitting.

EPSRC
Engineering and Physical Sciences Research Council

http://www.cl.cam.ac.uk/research/rainbow/projects/pupiltracking

UNIVERSITY OF CAMBRIDGE
Computer Laboratory
RAINBOW RESEARCH GROUP