Hybrid eye center localization using cascaded regression and robust circle fitting

Alex Levinshtein       Edmund Phung       Parham Aarabi
Customized Native iOS/Android/Web Apps
The need for iris detection

Virtual contact lenses try-on

Gaze tracking

Iris recognition
In-the-wild iris detection

Most methods use specialized devices

NIR images
Head mounts
Iris scanners

Our focus: real-time detection on mobile devices using a standard camera
Related Work

Hand-crafted methods

Ahuja et al. 2016
Fuhl et al. 2015, 2016
George and Routray 2016
Li et al. 2005
Skodras and Fakotakis 2015
Swirski et al. 2012
Timm and Barth 2011
Valenti and Gevers 2012
Wood and Bulling 2014

Learning-based methods

Markus et al. 2014
Tian et al. 2016
Zhou et al. 2015

Powerful regressors
Simple features
Simple regressors
Powerful features

More robust but not accurate

Accurate but not robust
Our contributions

1. Learning-based method using powerful regressors \textbf{and} powerful features

2. Increased accuracy by combining \#1 with robust circle fitting
Method Overview

Image + Facial landmarks

→

Cascade of regression forests

→

Accurate eye center and iris radii estimates after circle fitting

→

Rough eye center estimates
Rough eye center detection

Image + Facial landmarks

Cascade of regression forests

Accurate eye center and iris radii estimates after circle fitting

Rough eye center estimates
Cascade of gradient boosted regression forests

\[ S^{t+1} = S^t + r_t(I, S^t) \]

image I

pose \( S^t \)

\( r_t(I, S^t) \)

Shifts for left and right eyes

pose \( S^{t+1} \)
Cascade of gradient boosted regression forests

\[ r_0(I, S^0) = \{ \rightsquigarrow, \leftsquigarrow \} \]
Cascade of gradient boosted regression forests

\[ r_0(I, S^0) = \{ \rightarrow, \rightarrow \} \]

\[ r_1(I, S^1) = \{ \rightarrow, \rightarrow \} \]
Beyond raw pixels in $r_t(I,S_t)$

1. Use powerful features, instead of raw pixels
2. Anchor the features to current eye center estimates
Normalized HoG features for iris detection

Initial pose

\[ S^0 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \]

Features:
Pairwise differences of HoG

0.4 \cdot E_{\text{inter}}
Iris pose refinement by circle fitting

Image + Facial landmarks

Cascade of regression forests

Accurate eye center and iris radii estimates after circle fitting

Rough eye center estimates
Robust circle fitting

Fit iris using circle refinement

Start with \((a, b)\) set to regressor output and \(r\) set to 10% of interocular distance

\[
C_1 = \sum_{i=1}^{N} \left( \sqrt{(e_{ix} - a)^2 + (e_{iy} - b)^2} - r \right)^2
\]
Robust circle fitting

Issues:
1. Not robust to outliers
2. Only uses regressor output as initial guess

\[
C_1 = \sum_{i=1}^{N} \left( \sqrt{(e_{ix} - a)^2 + (e_{iy} - b)^2} - r \right)^2
\]
Robust circle fitting

\[ C_2 = w_1 \frac{1}{N} \sum_{i=1}^{N} \rho \left( \sqrt{(e_{ix} - a)^2 + (e_{iy} - b)^2} - r \right) + w_2 (a - a_0)^2 + w_2 (b - b_0)^2 + w_3 (r - r_{default})^2 \]
Robust circle fitting

1. Robust to outliers
2. Uses regressor output as initial guess and prior

\[ C_2 = w_1 \frac{1}{N} \sum_{i=1}^{N} \rho \left( \sqrt{(e_{ix} - a)^2 + (e_{iy} - b)^2} - r \right) + w_2 (a - a_0)^2 + w_2 (b - b_0)^2 + w_3 (r - r_{default})^2 \]
Results

Within the pupil iris

<table>
<thead>
<tr>
<th>Method</th>
<th>$e &lt; 0.025$</th>
<th>$e &lt; 0.05$</th>
<th>$e &lt; 0.1$</th>
<th>$e &lt; 0.25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BioID Dataset</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REG-R</td>
<td>68.13%</td>
<td>95.07%</td>
<td>99.59%</td>
<td>100%</td>
</tr>
<tr>
<td>REG</td>
<td>74.3%</td>
<td>95.27%</td>
<td>99.52%</td>
<td>100%</td>
</tr>
<tr>
<td>Timm [8]</td>
<td>38%*</td>
<td>82.5%</td>
<td>93.4%</td>
<td>98%</td>
</tr>
<tr>
<td>Valenti [9]</td>
<td>55%*</td>
<td>86.1%</td>
<td>91.7%</td>
<td>97.9%</td>
</tr>
<tr>
<td>Zhou [17]</td>
<td>50%*</td>
<td>93.8%</td>
<td>99.8%</td>
<td>99.9%</td>
</tr>
<tr>
<td>Ahuja [1]</td>
<td>NA</td>
<td>92.06%</td>
<td>97.96%</td>
<td>100%</td>
</tr>
<tr>
<td>Markuš [15]</td>
<td>61%*</td>
<td>89.9%</td>
<td>97.1%</td>
<td>99.7%</td>
</tr>
<tr>
<td>GI4E Dataset</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REG-R</td>
<td>88.34%</td>
<td>99.27%</td>
<td>99.92%</td>
<td>100%</td>
</tr>
<tr>
<td>REG</td>
<td>77.57%</td>
<td>99.03%</td>
<td>99.92%</td>
<td>100%</td>
</tr>
<tr>
<td>ELSE [3]</td>
<td>49.8%</td>
<td>91.5%</td>
<td>97.17%</td>
<td>99.51%</td>
</tr>
<tr>
<td>Anjith [4]</td>
<td>NA</td>
<td>89.28%</td>
<td>92.3%</td>
<td>NA</td>
</tr>
<tr>
<td>Talking Face Dataset</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REG-R</td>
<td>65.78%</td>
<td>95.68%</td>
<td>99.88%</td>
<td>99.98%</td>
</tr>
<tr>
<td>REG</td>
<td>18.7%</td>
<td>95.62%</td>
<td>99.88%</td>
<td>99.98%</td>
</tr>
<tr>
<td>ELSE [3]</td>
<td>59.26%</td>
<td>92%</td>
<td>98.98%</td>
<td>99.94%</td>
</tr>
<tr>
<td>Ahuja [1]</td>
<td>NA</td>
<td>94.78%</td>
<td>99%</td>
<td>99.42%</td>
</tr>
</tbody>
</table>

\[ e = \frac{1}{d} \max\{e_R, e_L\} \]
Results

Increased fine-level accuracy due to circle refinement

Majority of methods are not even evaluated at $e < 0.025$
Summary

1. Real-time iris detection system from standard images

2. Eye center localization using cascades of boosted regression trees with HoG features

3. Accurate iris localization using robust circle fitting

4. State-of-the-art performance on multiple datasets
Questions?