

000
001
002
003
004
005
006
007
008
009
010
011054
055
056
057
058
059
060
061
062
063
064
065
066
067
068
069
070
071
072
073
074
075
076
077
078
079

Rendering Photorealistic Training Images for Eye Tracking

Anonymous ICCV submission

Paper ID ****

Abstract

The ABSTRACT is to be in fully-justified italicized text, at the top of the left-hand column, below the author and affiliation information. Use the word “Abstract” as the title, in 12-point Times, boldface type, centered relative to the column, initially capitalized. The abstract is to be in 10-point, single-spaced type. Leave two blank lines after the Abstract, then begin the main text. Look at previous ICCV abstracts to get a feel for style and length.

1. Introduction

Machine learning approaches that leverage large amounts of image data are currently the best solutions to many problems in computer vision [cite]. However, capturing or collecting images can be extremely time consuming, especially for new areas of research without pre-existing datasets. Supervised learning approaches then require that the images are labelled. This annotation process can be expensive and tedious, and there is no guarantee the labels will be correct.

In this paper we describe our approach for generating photorealistic training data, and then present and evaluate two systems trained on SynthesEyes: an eye-region specific deformable model and an appearance-based gaze estimator. These systems are case studies that show how we leverage the degrees of control made available by rendering our training data to easily and quickly generate high quality training datasets.

2. Related work

2.1. Synthetic data

[7] – uses rendered videos of eyes to evaluate eye tracking algorithms.

[6] – relit 3d face scans to study the effect of illumination on automatic expression recognition.

[3] – train head pose estimator on only synthetic depth data.



(a) 3D eye model (b) Pupil dilation and iris color variation

Figure 1: Our realistic eye model is capable of expressing degrees of variability seen in real life.

2.2. Deformable eye model

[1] – trained a detailed deformable eye region model on in-the-wild images.

2.3. Gaze estimation

[8] – regression with features of 3d pupil centers and eye-contours (the eyelids) for gaze estimation. Use multiple cameras and IR lights.

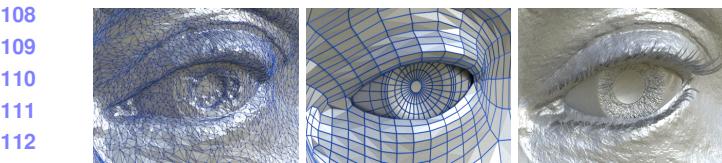
3. Synthetic data generation

In this section we first present our anatomically inspired CG eyeball model, and then explain our novel procedure for preparing a suite of 3D head scans for dynamic photorealistic labelled data generation. We then briefly describe how we use image-based lighting [2] to model a wide range of realistic lighting conditions, and finally discuss the details of our rendering setup.

3.1. Eye model

Eyeballs are complex organs comprised of multiple layers of tissue, each with different reflectance properties and levels of transparency. Fortunately, as realistic eyes are so important for many areas of CG, there is already a large body of previous work on modelling and rendering eyes Erroll: cite.

As shown in Figure 3c, our eye model consists of two parts. The outer part (red wireframe) approximates the eye’s overall shape with two spheres ($r_1 = 12\text{mm}$, $r_2 = 8\text{mm}$ [5]),



108
109
110
111
112
113
114
115
(a) 3D head scan data: (b) After retopology: (c) Detail is stored in
1.7 million polys 9 thousand polys a displacement map

116
117
118
Figure 3: Model preparation process

119
120
121
122
123
124
125
126
127
128
129
the latter representing the corneal bulge. To avoid a discontinuous seam between spheres, the meshes were joined and then smoothed. It is transparent, refractive ($n=1.376$), and partially reflective. The eye's bumpy surface variation is modelled by a displacement map generated with noise functions. The inner part (blue wireframe) is a flattened sphere with Lambertian material. The planar end represents the iris and pupil, and the rest represents the sclera – the white of the eye. There is a 0.5mm gap between the outer and inner parts which accounts for the thickness of the cornea. **Erroll: compare with recent Disney work**

130
131
132
133
134
135
136
137
Eyes exhibit variations in both shape (pupillary dilation) and texture (iris color and scleral veins). To model shape variation we use *shape keys* – a CG animation technique where different versions of a mesh are stored, modified, and interpolated between. We have shape keys for dilated and constricted pupils, as well as large and small irises to account for a small amount (10%) of iris size variation.

138
139
140
141
142
143
144
145
146
147
148
149
We vary the appearance of the eye by compositing textures in three separate layers: *i*) a *sclera* layer representing the tint of the sclera (white, pink, or yellow); *ii*) an *iris* layer with four photo-textures of different colored irises (amber, blue, brown, grey); and *iii*) a *veins* layer which varies between blood-shot and clear. We matched the sclera tint to each separate face model, but uniformly randomly varied iris color. Previous research on iris-synthesis **Erroll: cite** would have allowed continually different iris textures, but we decided this added complexity would not make a worthwhile improvement in overall appearance variation, especially when rendered at lower resolutions.

150 151 3.2. Preparing a suite of 3D eye-region models

152
153
Start with 3D head scan data – a dense mesh.
154
Remove scanned eye.
155
We want to be able to animate the model, so retopologize.
156
This also reduces render time and file sizes.

157
Insert eye-model.
158
Add eye lashes.
159
Add landmarks mesh.
160
Create face, eyelash, and landmark blend shapes for eye-
161
lids looking up and down.



162
163
164
165
166
167
168
169
Figure 4: Appearance variation from lighting is modelled
with poseable high-dynamic-range environment maps [2].

170 171 3.2.1 Eyelid motion

172
Vertical saccades are always accompanied by eyelid motion [4].

173 174 3.3. Lighting

175 176 4. Experiments

177 178 4.1. Deformable model

- 179
180
Evaluate eyelid landmark accuracy on LFW and M-
PIE data, compare against several state-of-the-art
CLM methods.
- 181
182
Evaluate eyelid and iris landmarks on hand-annotated
MPII data, compare against a baseline method: major-
ity vote for iris position.

183
184
185
186
187
(Maybe) Plot landmark accuracy on LFW against num-
ber of training participants. Show that even with just a few
participants (e.g. 4) we get good results for eyelid positions
compared to state-of-the-art face trackers.

188 189 4.2. Gaze estimation

190
191
192
193
194
We render a targeted dataset that matches MPII's gaze
and pose distribution, with added 3D laptop screen emitting
light. This shows how we can target specific scenarios like
laptop-based gaze estimation, and render a suitable dataset
within a day rather than take 3 months of data collection.

195
196
197
198
199
Using Xucong's CNN system, we train on targeted ver-
sion of SynthesEyes, test on MPII. Show results are better
than training on UT and testing on MPII. This shows that
the range of lighting in SynthesEyes is important for better
results.

200 201 5. Conclusion

202 203 References

- [1] J. Alabert-i Medina, B. Qu, and S. Zafeiriou. Statistically learned deformable eye models. In *Computer Vision-ECCV 2014 Workshops*, pages 285–295. Springer, 2014. 1
- [2] P.Debevec. Image-based lighting. *IEEE Computer Graphics and Applications*, 22(2):26–34, 2002. 1, 2
- [3] G. Fanelli, J. Gall, and L. Van Gool. Real time head pose es-
timation with random regression forests. In *Computer Vision* 210
211
212
213
214
215



Figure 2: Our suite of female and male head models for rendering.

and Pattern Recognition (CVPR), 2011 IEEE Conference on,
pages 617–624. IEEE, 2011. 1

- [4] S. Liversedge, I. Gilchrist, and S. Everling. *The Oxford handbook of eye movements*. Oxford University Press, 2011. 2
- [5] K. Ruhland, S. Andrist, J. Badler, C. Peters, N. Badler, M. Gleicher, B. Mutlu, and R. McDonnell. Look me in the eyes: A survey of eye and gaze animation for virtual agents and artificial systems. In *Eurographics State-of-the-Art Report*, pages 69–91, 2014. 1
- [6] G. Stratou, A. Ghosh, P. Debevec, and L. Morency. Effect of illumination on automatic expression recognition: a novel 3d relightable facial database. In *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*, pages 611–618. IEEE, 2011. 1
- [7] L. Świrski and N. Dodgson. Rendering synthetic ground truth images for eye tracker evaluation. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, pages 219–222. ACM, 2014. 1
- [8] C. Xiong, L. Huang, and C. Liu. Gaze estimation based on 3d face structure and pupil centers. In *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pages 1156–1161. IEEE, 2014. 1

216
217
218
219
220
221
222

270
271
272
273
274
275
276

223
224
225

277
278
279

226 and Pattern Recognition (CVPR), 2011 IEEE Conference on,
227 pages 617–624. IEEE, 2011. 1

280
281

- [4] S. Liversedge, I. Gilchrist, and S. Everling. *The Oxford handbook of eye movements*. Oxford University Press, 2011. 2
- [5] K. Ruhland, S. Andrist, J. Badler, C. Peters, N. Badler, M. Gleicher, B. Mutlu, and R. McDonnell. Look me in the eyes: A survey of eye and gaze animation for virtual agents and artificial systems. In *Eurographics State-of-the-Art Report*, pages 69–91, 2014. 1
- [6] G. Stratou, A. Ghosh, P. Debevec, and L. Morency. Effect of illumination on automatic expression recognition: a novel 3d relightable facial database. In *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*, pages 611–618. IEEE, 2011. 1
- [7] L. Świrski and N. Dodgson. Rendering synthetic ground truth images for eye tracker evaluation. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, pages 219–222. ACM, 2014. 1
- [8] C. Xiong, L. Huang, and C. Liu. Gaze estimation based on 3d face structure and pupil centers. In *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pages 1156–1161. IEEE, 2014. 1

282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323

247
248
249
250
251
252
253
254
255
256
257
258
259
260

301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320

261
262
263
264
265
266
267
268
269

315
316
317
318
319
320
321
322
323