Vision, Video and Graphics (2005)
E. Trucco, M. Chantler (Editors)

Correspondenceless Stereo for 3-D Iris Location

Tom Anderson, Emanuele Trucco and Marco Razeto

Vision, Signal and Image Processing Group
School of Engineering and Physical Sciences, Heriot Watt University, Edinburgh, UK
{twa1,e.trucco}@hw.ac.uk

Abstract

We present a correspondenceless stereo system locating the circular contour of an iris (limbus) in space, and therefore its 3-D plane. We avoid correspondence search completely by intersecting a bundle of conjugate epipolar lines with the elliptical images of the limbus in the stereo pair of images, which gives correspondences directly. The ellipses are located by active ellipse fitting. An efficient simulated annealing implementation achieves reliable iris location with uncontrolled illumination and eye or head movements. Tests with ground-truthed 3-D setups as well as real eye images indicate very good accuracy.

1. Introduction and motivation

We present a stereo system locating an iris in space from images of a single eye, but avoiding correspondence search altogether. The system works in normal room illumination and unconstrained eye movements.

Iris location and gaze detection systems are an essential component of many systems requiring instantaneous location of the viewpoint of an observer (e.g., see Isgró et al. [ITKS04] for image-based communications), and play an important role for security and safety [J.D04, LTYD03, CDAM00, SL04] as well as medical applications [DTCT02, KMRW03, NRJ*99].

Existing techniques for iris location are *invasive* or *non-invasive*. Invasive techniques involve devices applied to the subject, e.g., electrodes, contact lenses, head-mounted photodiodes or cameras [A.T03]. Non-invasive techniques do not use such devices, but rely very often on special illumination to highlight relevant eye characteristics, e.g., reflections of infrared light off the cornea-lens boundaries (Purkinje images [SL04, CDAM00, ZJFL02]). We do not use either controlled illumination or external devices, but rely solely on a stereo pair of images.

The assumption of uncontrolled llumination, eye movements and iris appearance prevents us from using well-established techniques in iris recognition for people identification. Ma et al. [LTYD03] use only high quality, unoccluded frontal images and detect the limbus with a combination of edge detection and Hough transform. The integro-differential operator used by Daugman [J.D93, J.D04] in the iris recognition context proves very robust, but is restricted to frontal images where the limbus appears circular. To account for arbitrary orientation of the iris plane we adopt a detection technique based on active ellipses.

Active ellipses and *active contours* have been widely studied in computer vision [IB98, MMM00, LMJM01, SWAC02] and in medical applications [KMRW03, DTCT02, JJJ97]. We detect the iris without relying on context, instead of using a complete eye template [AA92, CC03, BCT02], because our images do not include necessarily other facial feature, or these may not be detected reliably. Specific problems are iris occlusion, unwanted features (e.g., corneal reflections) and varying image quality (e.g., blur, skin type, race). Occlusion is the biggest problem, since part of the iris is almost invariably covered by the eyelids.

To combat the above, we use a robust active ellipse location algorithm that recovers the limbus even when limited arcs are visible. The algorithms maximises a criterion comparing intensity variations across the perimeter of a candidate ellipse with an a-priori, multi-scale model synthesized from extensive observations. Robustness is provided by an efficient simulated annealing (SA henceforth) search with parameter values optimised for our applications via extensive, systematic testing on a set of 327 ground-truthed images. Non-stochastic algorithms tried (gradient descent, Nelder-Mead) failed to reach the correct minimum with re-

motely comparable consistency, and, unlike stochastic optimizers, were extremely sensitive to initialisation.

Algorithms have been reported to locate the iris and/or its orientation in space (necessary for gaze estimation), for instance from an image capturing both eyes simultaneously [WS00] or with stereo systems [BF03, NMRZ00], with typical correspondence search. We locate the iris circle in space using stereo, but without correspondence. This is possible as we are looking for two corresponding ellipses (projection of the same 3-D circle); once we have located a candidate pair, the intersections with conjugate epipolar lines provide point correspondences directly. This idea is akin to Cham and Cipolla's stereo coupled contours [CC97], and more generally linked to the coupling of matching and reconstruction [KA03]. 3-D points are calculated from the correspondences by calibrated triangulation. The best-fit 3-D circle to the reconstructed points is the sought iris contour in space. We plan to incorporate a closed-form reconstruction of 3-D conics from two image projections [Qua96, SZ00] at a later stage.

This paper makes two main contributions: a model-based stereo system avoiding correspondence altogether, and a robust iris (limbus) detection system via active ellipse fit and efficient stochastic search. We do not describe the target application in detail for industrial confidentiality reasons.

In the remainder of the paper, Section 2 describes the iris location, in particular the shape and intensity-based cost function and the optimisation scheme adopted for robust ellipse detection. Section 3 describes our correspondenceless stereo matching and associated reconstruction of the iris plane in space. Section 4 summarises our extensive tests assessing the accuracy of the iris detection module, and tests with the complete iris location system, first with a controlled, ground-truthed setup enabling accuracy measurements, then with real eye images. Section 5 sketches some conclusions and future work.

2. Robust limbus detection via active ellipses

2.1. Input, output and preprocessing

The input is a monochrome image of a single eye; the output is an ellipse tracing the contour of the iris, illustrated in Figure 1. We suppress corneal reflections, motion blur and other artefacts, introducing distracting, strong contours, with a 10×10 median filter. This relatively large mask affects texture but the algorithm only boundary information, well preserved by median filtering.

2.2. Modelling the iris contour for optimization

We find the limbus via an optimization in the parameter space of an active ellipse model. The result is the ellipse which best follows a predicted, multi-scale model of intensity changes.

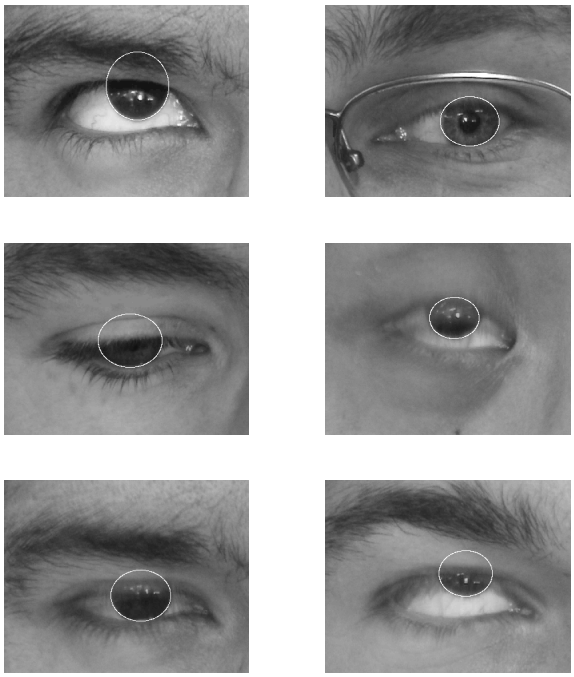

Figure 1: *Examples of iris detection results; notice occlusions (some heavy), and uncontrolled illumination and pose.*

The visible portion of the limbus is characterized by a noisy light (sclera) to dark (iris) intensity transition. In our application and imaging set-up, the spatial extent of such transition is between 3 and 12 pixels approximately. We therefore model the transition with two Petrou-Kittler ramp edges [MJ91] at two different spatial scales.

The ellipse is parametrized by its axes and centre coordinates. The axes are assumed aligned with the image axes, as tilt assumed limited (and therefore ellipse inclination negligeable) in our images. The cost function extracts intensity profiles along 30 normals to the candidate ellipse, distributed uniformly (Figure 2). These profiles are, ideally, convolved with two optimal ramp detection filter masks (Petrou and Kittler [MJ91], see below) at two different spatial scales. In practice, we are interested only in the filter output at the centre of the normal segments (i.e., on the ellipse perimeter), so we compute only *one* filtered value per segment. Filtered values are summed over all normals and over both filter sizes to obtain the criterion to optimize, c:

$$c = -\sum_{i=1}^{N}\left(\int_{-w}^{w} S_i(x) f_1(x)\,dx + \int_{-w}^{w} S_i(x) f_2(x)\,dx\right), \qquad (1)$$

where N is the number of control points, S_i is the intensity profile extracted at the control point i and f_1 and f_2 are the

filters at the two different scales. Examples of Petrou-Kittler filters are given in Figure 2; their analytical form is

$$h(x) = e^{Ax}[L_1 \sin(Ax) + L_2 \cos(Ax)] + \\ + e^{-Ax}[L_3 \sin(Ax) + L_4 \cos(Ax)] + L_5 x + L_6 e^{sx} + L_7,$$

for $-w \leq x \leq 0$, where w is the filter half-width. The values of A and $L_1 \ldots L_7$ are tabulated for the two target spatial scales. The parameter s represents the inverse of the transition length. We can easily tune the filter for different values of s by rescaling A and w. These filters are antisymmetric and therefore unaffected by uniform changes of illumination. Extensive testing identified masks optimal for ramps of width 4 and 10 pixels as responding well to typical limbus edges in our target images, and poorly to most transitions related to non-iris features such as eyebrows or eyelashes.

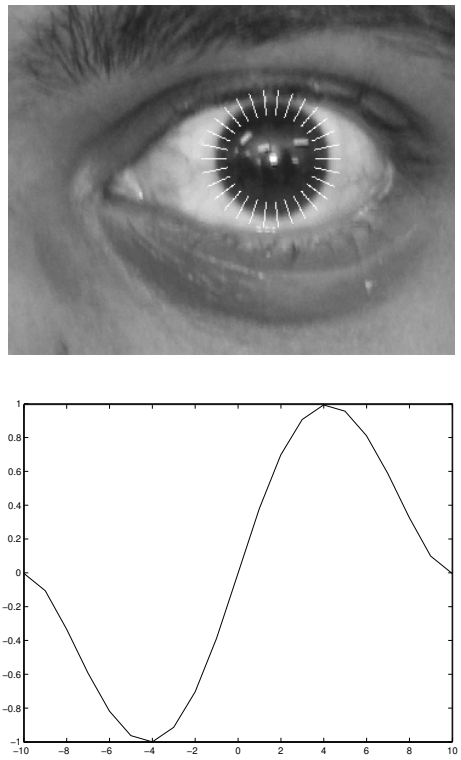

Figure 2: *Top: segments normal to the candidate ellipse (here, overlapping the iris). Bottom: example of Petrou-Kittler optimal ramp detector (ramp extent is 10).*

2.3. Optimization scheme

Once established that deterministic search would not be enough, we assessed the performance of various non-deterministic optimizers for our problem. We considered standard SA, two SA variations (great deluge [DS90] and thresholded annealing [Due93]), and Girosi-Caprile [CG90]. We recorded estimation errors in the four ellipse parameters considered over extensive ranges of variation of the algorithms' parameters, taking care to keep algorithms working in comparable conditions. This work is detailed in [S.R04]. The result indicated standard SA as marginal winner over thresholded annealing.

We summarise here our operational choices only, and refer interested readers to Slamon et al. [PPR02] for a full treatment of SA and its practicalities, and to [RT05] for details of our implementation. The active ellipse is parameterized by a, b (semiaxes), O_x, O_y (centre co-ordinates), forming a 4-D state vector, S. Small inclinations are possible (and indeed included in our tests) as the head is not constrained, but do not upset the fit enough to justify the cost of adding a fifth element to the state vector. This is initialised at the image centre with a default size. The number of ellipses tested is progressively reduced with temperature, in our tests from $T_{start} = 500$ to $T_{end} = 1$. These values were decided by sampling the cost function over several images and calculating the relative acceptance ratio, whose desirable value at high temperature is around 50%.

New candidate values are generated for each parameter from a Gaussian distribution centered in the previous value, with standard deviation $\sigma_{new} = R\sigma_{old}$, much in the spirit of Ingber's [L.I89] fast annealing. R controls the search range, starting from 2 pixels for ellipse centre and 1 pixel for axes lengths and decreasing with an indipendent annealing schedule (see below).

The acceptance rule for new states is the standard Metropolis rule: candidate states bringing an energy increase are accepted with a probability $e^{-\Delta/T}$ (notice the dependence on T). The annealing schedule affects the move class via the range parameter R:

$$R_{new} = (\frac{1}{\sqrt{t+1}} + 0.3) R_{old}$$

where $t4$ is the annealing iteration index (time).

3. Stereo without correspondences

Correspondence search can be avoided as corresponding known curves are located in both images, and the epipolar geometry is known. The intersection of conjugate epipolar lines with such curves provide corresponding points, and, through reconstruction, 3-D points in space. This is illustrated in Figure 3. Notice that, in our imaging assumptions, a single ellipse is located in each image, so that no ambiguity exists (although location errors are possible). A model of the 3-D curve (here, a circle) can then be fitted to the reconstructed 3-D points.

We obtain the epipolar geometry from full calibration, but of course weak calibration (only image correspondences

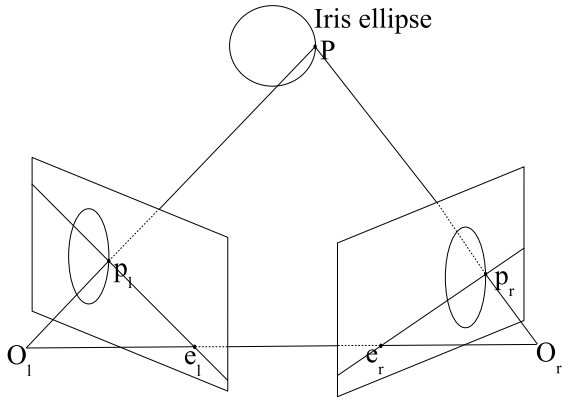

Figure 3: *Corresponding points as intersections of conjugate epipolar lines with the left and right projections (ellipses) of a 3-D circle.*

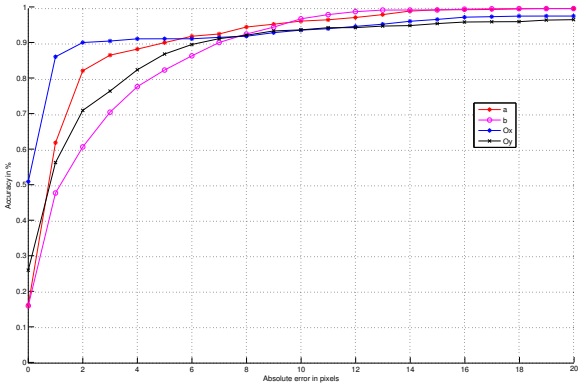

Figure 4: *Monte-Carlo estimates of error probability of absolute error (in pixels) for the four ellipse parameters. X axis: absolute error in pixel. Y axis: relative frequencies (probability estimates).*

known) would suit as well, at least for estimating the orientation of the iris plane in space [TV98]. Notice that the size of the human iris appears to be very stable across individuals and even races [A.T03], so that reasonable distance estimates could be easily achieved even with weakly calibrated cameras.

Figure 6 shows two pairs of typical images, the detected irises, and the bundles of conjugate epipolar lines used for correspondence. The epipolar bundle must be chosen in such a way to guarantee accurate intersections in both images, i.e., the epipolar lines must be as normal to the ellipse as possible at the intersection points. To this purpose we choose 20 points on the left-image ellipse avoiding the top and bottom arcs, where epipolar line may approach the ellipse tangent. The points are spaced by 10° intervals along the ellipse, and grouped in two sets symmetric with respect to the vertical ellipse axis.

The 3-D plane best fitting the reconstructed points is found by linear least squares via singular value decomposition. Robust fitting is unnecessary as surely no outliers are present: correspondences are drawn from pre-fitted parametric curves.

4. Experimental results

4.1. Iris detection accuracy

To test the accuracy of iris detection, we used a database of 327 monochrome test images with various amounts of iris occlusion and blur, various gaze directions, different skin colours and eye shapes, with and without spectacles. The images were 350×270 and captured by a digital camera or a digital camcoder with uncontrolled room lighting. Ground truth was established manually by tracing ellipses following the limbus in each image. We performed 50 runs on each image ($50 \times 327 = 16,350$ runs). The ellipse is initialised always at the image centre, with semiaxes of 40 pixels each (the initial position is close to immaterial for SA). We computed the difference between estimates of ellipse parameters and the corresponding ground truth values. Performance analyses for each parameter are reported elsewhere [RT05]; here, we summarize performance in Figure 4. This graph shows, for each ellipse parameter, Monte-Carlo estimates of the cumulative probability of a given error value in pixels (in actual fact, relative frequencies). The graph is obtained by integrating the error histograms plotted for each parameters. For instance, 91.5% of the O_x histogram falls within a 5-pixel tolerance interval, suggesting an *approximate* probability of 91.5% for this accuracy level of the horizontal component. For O_y, this figure is 88%. A 5-pixel accuracy on *all* parameter is likely to be suitable for several applications.

4.2. Correspondenceless stereo accuracy

We present two series of tests, with controlled and real setups respectively. All results were achieved with a MATLAB implementation running on a Pentium III PC under Windows. Monochrome, PAL-resolution stereo pairs were acquired with PULNIX PEC3010 cameras and a Matrox Meteor II frame grabber. The stereo pair was calibrated using Tsai's classic procedure [TV98].

Controlled tests. To establish quantitative ground truth against which to compare estimates of iris plane orientation, we fixed a picture of a real iris onto a planar support. This support was rotated through an interval of 15 degrees around a vertical axis in steps of 1 degree, measured manually using a protractor. The interval was centered around the head-on direction (iris normal along the Z axis, pointing towards the cameras). For each angle, we estimated the orientation or the iris plane. The cameras were calibrated so that the axis of rotation was the X axis of the world reference frame, allow-

ing consistent comparisons of estimates and ground truth. The interocular distance was 90mm, the focal lengths 12.48 and 10.74mm, and the stand-off distance (from left camera) about 200mm. Figure 5 shows the angular error on the XZ

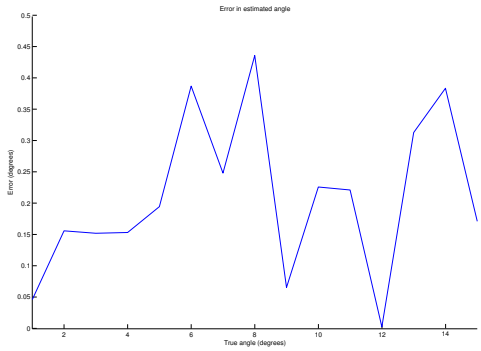

Figure 5: *Estimated angular difference (degrees) between projections onto the XZ (ground) plane of estimated and true normals to the iris plane.*

(ground) plane for each orientation . This error is defined as the angular difference (in degrees) between the normals to the true and estimated iris plane after projection on the XZ plane. The mean is 0.21°, the standard deviation 0.13°, both below the accuracy with which we could measure ground truth quantities. The full error, i.e., the angular difference between full (not projected) normals, is larger, in part because our manual positioning system did not guarantee repeatable orientations nor perfectly vertical iris planes, in part because estimated normals did include a small Y component. The mean of the full error was 1.5° and the standard deviation 0.4°, ostensibly still very good results.

Real-eye tests. The same camera setup as before was used to acquire stereo images of real eyes. Examples of images with a superimposed bundle of epipolar lines intersecting the detected ellipse are shown in Figure 6. together with the ellipse arcs fitted in 3-D space. The mean deviation from best-fit planes was 0.1mm, with average standard deviation 0.13mm, and maximum deviation of less than 1mm, suggesting an accurately planar reconstruction. We cannot draw conclusions on absolute orientation accuracy, for which we rely on the controlled tests.

5. Conclusions

We have presented a correspondenceless stereo system locating an iris in 3-D space. We avoid correspondence search completely by intersecting a bundle of corresponding epipolar lines with the ellipses images of the iris in the two images. The intersections give corresponding points for free. The elliptical iris contour is located by an active ellipse fit based on an efficient simulated annealing implementation,

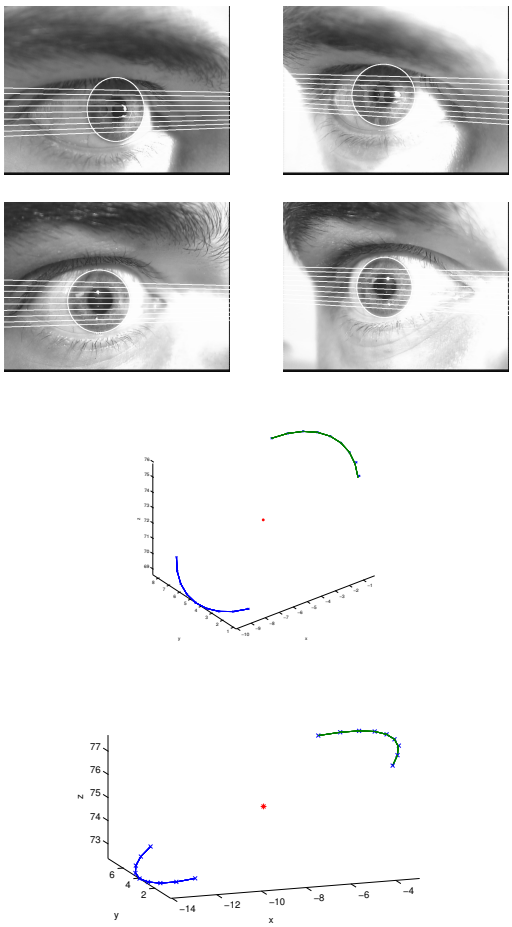

Figure 6: *First and second rows: The located limbus and the bundle of conjugate epipolar lines in two real stereo pairs. Bottom: 3-D reconstruction of circle arcs for first and second pair, respectively.*

which proved very accurate in extensive tests. Results with controlled 3-D set-ups providing ground truth for iris plane orientation indicate low error levels. Tests with images of real eyes suggest very good planar fits. We conclude that the overall reliability of the system for gaze direction estimation is very good.

Future works will address stereo gaze tracking, the use of epipolar constraints within the iris location algorithms, and incorporating closed-form reconstruction of the 3-D circle from the image projections.

References

[AA92] A.BLAKE, A.L.YUILLE: *Active Vision.* MIT Press, 1992.

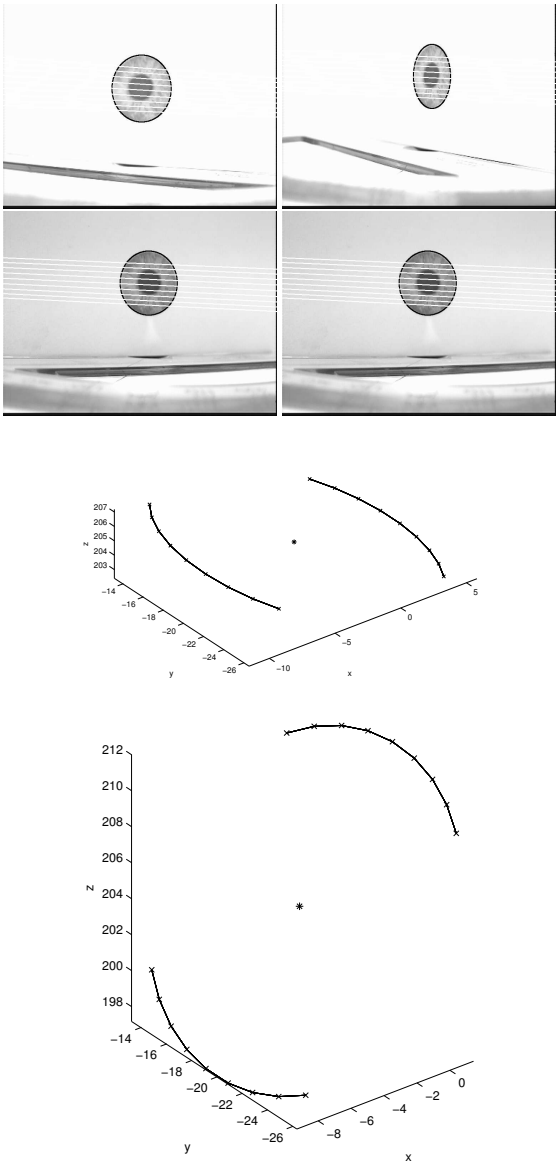

Figure 7: *First and second rows: The located limbus and the bundle of conjugate epipolar lines in two artificial eye image pairs. Bottom: 3-D reconstruction of circle arcs for first and second pair, respectively.*

[A.T03] A.T.DUCHOWSKY: *Eye tracking methodology*. Springer, 2003.

[BCT02] BETTINGER F., COOTES T., TAYLOR C.: Modelling facial behaviours. In *Proc. BMVC'02* (2002), vol. 2, pp. 797–806.

[BF03] BEYMER D., FLICKNER J.: Eye gaze tracking using an active stereo head. In *IEEE Int Conf on Comp Vis and Patt Rec (CVPR)* (2003), vol. 2, pp. 18–20.

[CC97] CHAM T.-J., CIPOLLA R.: Stereo coupled active contours. In *IEEE Int Conf on Comp Vis Patt Rec (CVPR)* (1997), pp. 1094–1099.

[CC03] CRISTINACCE D., COOTES T.: Facial feature detection using ADABOOST with shape constraints. In *Proc BMVC'03* (2003), vol. 1, pp. 231–240.

[CDAM00] C.H.MORIMOTO, D.KOONS, A.AMIR, M.FLICKNER: Pupil detection and tracking using multiple light sources. *Image and Vision Computing* (2000).

[CG90] CAPRILE B., GIROSI F.: *A nondeterministic optimisation algorithm*. Master's thesis, AI Memo 1254, Artif Intell Lab, MIT, 1990.

[DS90] DUECK G., SCHEUER T.: Threshold accepting: a general purpose optimization algorithm appearing superior to simulated annealing. *Journ of Computational Physics 90* (1990), 161–175.

[DTCT02] DAVIS R. H., TWINING C., COOTES T., TAYLOR C.: A minimum description length approach to statistical shape modelling. *IEEE Trans Medical Imaging 21* (2002), 535–527.

[Due93] DUECK G.: New optimization heuristics, the great deluge algorithm and the record-to-record-travel. *Journ of Computational Physics 104* (1993), 86–92.

[IB98] ISARD M., BLAKE A.: *Active Contours*. Springer, 1998.

[ITKS04] ISGRÒ F., TRUCCO E., KAUFF P., SCHREER O.: 3-d image processing in the future of immersive media. *IEEE Trans on Circuits and Systems for Video Technology 14*, 3 (2004), 288–303.

[J.D93] J.DAUGMAN: High confidence visual recognition of persons by a test of statistical independence. *IEEE Trans PAMI 15*, 11 (1993).

[J.D04] J.DAUGMAN: How iris recognition works. *IEEE Trans Circuits and systems for video technology 14*, 1 (2004).

[JJJ97] J.P.IVINS, J.PORRILL, J.P.FRISBY: A deformable model of the human iris driven by non-linear least-squares minimisation. In *Proc. of IPA'97* (1997), vol. 1, pp. 234 – 238.

[KA03] KAHL F., AUGUST J.: Multiview reconstruction of space curves. In *IEEE Int Conf on Comp Vis (ICCV)* (2003), vol. 2, pp. 1017–1024.

[KMRW03] KERRIGAN S., MCKENNA S., RICKETTS I. W., WIDGEROWITZ C.: Analysis of total hip replacements using active ellipses. In *Proc. Medical Images Understanding and Analysis (MIUA)* (2003).

[L.I89] L.INGBER: Very fast simulated re-annealing. *Mathematical and computer modelling 12* (1989), 967–973.

[LMJM01] LEROY B., MEDIONI G. G., JOHNSON E., MATTHIES L.: Crater detection for autonomous landing on asteroids. *Image and Vision Computing 19*, 11 (2001), 787–792.

[LTYD03] L.MA, T.TAN, Y.WANG, D.ZHANG: Personal identification based on iris texture analysis. *IEEE Trans PAMI 25*, 12 (2003).

[MJ91] M.PETROU, J.KITTLER: Optimal edge detectors for ramp edges. *IEEE Trans. PAMI 13* (May 1991).

[MMM00] M.VINCZE, M.AYROMLOU, M.ZILLICH: Fast tracking of ellipses using edge-projected integration of cues. In *Int Conf on Pattern Recognition (ICPR'00)* (September 2000).

[NMRZ00] NEWMAN R., MATSUMOTO Y., ROUGEAUX S., ZELINSKY A.: Real-time stereo tracking for head pose and gaze estimation. In *4th Int Conf on Autom Face and Gesture Recogn* (2000), pp. 122–128.

[NRJ*99] N.RITTER, R.OWENS, J.COOPER, R.EIKELBOOM, SAARLOS P.: Registration of stereo and temporal images of the retina. *IEEE Trans. on Medical Imaging 18*, 5 (May 1999).

[PPR02] P.SALAMON, P.SIBANI, R.FROST: *Facts, Conjectures and Improvements for Simulated Annealing*. SIAM, 2002.

[Qua96] QUAN L.: Conic reconstruction and correspondence from two views. *IEEE Trans Patt Anal Mach Intell 18*, 2 (1996), 1–13.

[RT05] RAZETO M., TRUCCO E.: Robust iris location in close-up images of the eye. *Pattern Analysis and Applications* (2005), to appear.

[SL04] SHIH S.-W., LIU J.: A novel approach to 3-d gaze tracking using stereo cameras. *IEEE SMC (Part B) 34*, 1 (2004), 234–245.

[S.R04] S.RICHARD: *Comparative experimental assessment of four optimisation algorithms applied to iris location*. Master's thesis, ECE/EPS, Heriot-Watt University, 2004.

[SWAC02] S.RAMADAN, W.ABD-ALMAGEED, C.E.SMITH: Eye tracking using active deformable models. In *Proc. ICVGIP'02* (2002).

[SZ00] SCHMID C., ZISSERMAN A.: The geometry and matching of curves in multiple views. *Int Journ of Comp Vis 40*, 3 (2000), 199–233.

[TV98] TRUCCO E., VERRI A.: *Introductory Techniques for 3-D Computer Vision*. Prentice-Hall, 1998.

[WS00] WANG J. G., SUNG E.: Gaze determination via images of the iris. In *British Mach Vis Conf (BMVC)* (2000), vol. 1, pp. 132–141.

[ZJFL02] ZHU Z., JI Q., FUJIMURA K., LEE K.: Combining kalman filtering and mean shift for real-time eye tracking under active IR illumination. In *Int Conf on Patt Rec (ICPR)* (2002), pp. 318–321.