

The Image Ray Transform for Structural Feature Detection

Alastair H. Cummings, Mark S. Nixon, John N. Carter

*School of Electronics and Computer Science, University of Southampton, Southampton,
SO17 1BJ, United Kingdom*

Abstract

The use of analogies to physical phenomena is an exciting paradigm in computer vision that allows unorthodox approaches to feature extraction, creating new techniques with unique properties. A technique known as the “image ray transform” has been developed based upon an analogy to the propagation of light as rays. The transform analogises an image to a set of glass blocks with refractive index linked to pixel properties and then casts a large number of rays through the image. The course of these rays is accumulated into an output image. The technique can successfully extract tubular and circular features and we show successful circle detection, ear biometrics and retinal vessel extraction. The transform has also been extended through the use of multiple rays arranged as a beam to increase robustness to noise, and we show quantitative results for fully automatic ear recognition, achieving 95.2% rank one recognition across 63 subjects.

Keywords: vision, feature extraction, circle detection, biometrics, medical imaging

Email addresses: ahc08r@ecs.soton.ac.uk (Alastair H. Cummings),
msn@ecs.soton.ac.uk (Mark S. Nixon), jnc@ecs.soton.ac.uk (John N. Carter)

URL: <http://users.ecs.soton.ac.uk/ahc08r/> (Alastair H. Cummings)

Preprint submitted to Pattern Recognition Letters

August 23, 2011

1. Introduction

The detection of structural features within images is a challenging but vital problem in computer vision. Applications for its use are wide, ranging from machine vision to biometrics and medical imaging. We present the image ray transform (IRT), a technique based upon an analogy to light rays which has successfully been used for the detection of different structures within images.

There are a number of ways to approach the creation of techniques for computer vision. One paradigm that produces unique methods, results and advantages is the use of analogies to physical phenomena as a foundation (Nixon et al., 2011). Few techniques exist based upon analogies to light, but those that do often use the Eikonal equation. The Eikonal equation describes the time a ray takes to travel from the boundary of an anisotropic medium to a point within it. Examples of its use are for image half-toning (Pnueli and Bruckstein, 1994) and in a watershed segmentation method (Meyer and Maragos, 1999), although primarily as a distance measure rather than taking fuller advantage of the light analogy. The image ray transform is based upon an analogy to light rays, but is not a detailed simulation of light propagation; aspects of the analogy that enhance our technique are kept, but not those that weaken it.

This paper describes this novel technique employing the principles of ray optics to emphasise image features, particularly of a structural nature such as tubes and circles. In section 2 we describe the image ray transform, including the laws of optics on which the transform is based, the mechanics through which it works, as well as discussing properties and parameter selection and detailing an extension involving using a beam of rays to make

the transform robust to noise. Section 3 summarises how the transform can detect structural features, looking at circles as well as tubular features found in retinal images and ear biometrics, and also presenting new recognition results. Finally conclusions are drawn and some possible future applications for the transform are identified in section 4.

2. The Image Ray Transform

The image ray transform is a novel technique for the enhancement of structural features within images. It employs the principles of ray optics to highlight tubular and circular features within images. Although the transform is based on an analogy to light, the method can be modified to enhance feature detection or improve performance whilst still maintaining the advantages of the analogical formulation.

2.1. *Laws of Optics*

The propagation of light can be simulated through the use of rays, a model that takes specific regard for the direction of travel but ignores more complex properties like diffraction. Light rays travel through a medium, and at boundaries with other media may change in direction through either refraction and reflection, or total internal reflection. In the physical world light will often split, some part reflected and some refracted but in our model we consider only the possibility of refraction or total internal reflection.

When light crosses a boundary between two media, it does so at an angle of θ_I to the normal of the media boundary, as shown in figure 1. Reflection is when the light remains in the first media and continues travelling on the other side of the normal, where $\theta_I = \theta_L$. Refraction is when the

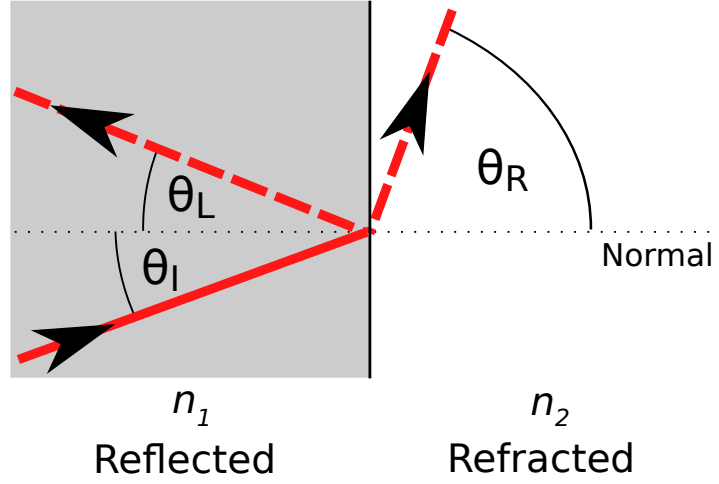


Figure 1: Refraction and reflection of light at a boundary of two media m_1 and m_2 .

light continues past the boundary into the second media, but has its angle modified so that it travels in direction θ_R to the normal instead. Which of these occur, as well as the angle of refraction, depends upon the refractive indices of the boundary media, n_1 and n_2 . Refractive indices are the ratio of the speed of light in a vacuum to the speed of light in a medium, and in nature $1 \leq n \lesssim 4$.

The angle of refraction θ_R is calculated from Snell's Law as:

$$\frac{\sin \theta_I}{\sin \theta_R} = \frac{n_2}{n_1} \quad (1)$$

The result of this is that if $n_1 > n_2$, then $\theta_R < \theta_I$, and light bends towards the normal. If $n_2 > n_1$ then we must take into account whether θ_I is above the critical angle, that is, the angle that would make $\theta_R = 90^\circ$ and so impossible:

$$\theta_C = \sin^{-1} \left(\frac{n_2}{n_1} \right) \quad (2)$$

If $\theta_I > \theta_C$ light must reflect through total internal reflection, if not then

$\theta_R > \theta_I$ and light bends away from the normal. These rules form the basis of the image ray transform.

2.2. Image Ray Transform Mechanics

The IRT operates by casting rays through an image. The transformed image is generated by recording the course of each ray, which results in enhanced structural features. We analogue the image as an array of two dimensional glass blocks, each representing a pixel in the image. Each block's refractive index is derived from the intensity of its corresponding pixel, through either a linear relationship, with a parameter controlling the maximum refractive index (n_{\max}). For a given intensity, i , the refractive index, n_i , is

$$n_i = 1 + \left(\frac{i}{255} \right) \cdot (n_{\max} - 1). \quad (3)$$

or exponential relationship, with a parameter k controlling the growth with intensity:

$$n_i = e^{\frac{i}{k}}. \quad (4)$$

Within this array of blocks, a ray is created with a random (subpixel) position (x and y) and an initial direction (ϕ) from a uniform distribution(U). A unit vector representing the ray's direction, \mathbf{V} , is calculated from ϕ . The ray is initialised at position vector \mathbf{p} :

$$\mathbf{p}^{<0>} = (x, y)^T \quad (5)$$

and at iteration t :

$$\mathbf{p}^{<t>} = \mathbf{p}^{<t-1>} + \mathbf{V} \quad (6)$$

When a ray enters a block for the first time, the accumulator matrix (\mathbf{A}) entry for that block is incremented by 1. Each individual ray increments a

pixel once in order to prevent small loops in the ray’s path causing small clusters of pixels to be repeatedly incremented, which are more likely to be caused by noise than a desired structural feature.

When the ray crosses a block boundary where media on either side have different refractive indices, a vector formation of the laws described in section 2.1 must be used to calculate a new direction. Through use of the direction of the ray \mathbf{V} , the normal \mathbf{N} , and the refractive indices of the two media n_1 and n_2 the new direction can be calculated. The equations for calculating the direction for refraction \mathbf{R}_r , direction for reflection \mathbf{R}_l from \mathbf{V} and \mathbf{N} and choosing between them can be found in Hill (2000). The ray is followed in this manner until it has undergone d reflections or refractions, travelled a distance l or the ray exits the image. This is repeated for N rays and the transformed image is found from the accumulator \mathbf{A} .

Whilst the method is strongly based upon an analogy to light rays, there are two changes we make in order to improve the result of the transform without compromising the analogy. The normals (\mathbf{N}) that are used to calculate the new direction of the ray are not set to horizontal or vertical block boundaries as would be expected, but rather the normal of the edge direction found by the Sobel operator at that point as this is more representative of the information within the image. Secondly, we only perform refraction when moving from higher to lower refractive indices as this tends to improve the quality of the transform result, and enhance structural features more strongly.

2.2.1. Parameter Selection

As with most techniques, parameter selection is an important step in ensuring high quality results. Whilst the ray transform has a number of parameters, it has been designed so that many of them are derived from image information, or that static values suggested here are appropriate. Parameter selection is primarily concerned with balancing noise and the speed of the transform. As the transform is non-deterministic due to the random initialisation of rays, a sufficient distribution of rays to cover pertinent features and limit noise is necessary. Conversely, every step of following a ray has a computational cost, and that must be minimised at the same time.

The most important parameters are those which define the relationship between refractive index and intensity, n_{\max} or k , depending on whether equation 3 or 4 is used. The difference between refractive indices changes with this value, and hence the amount that rays conform to structural features increases. Natural values of n_{\max} (below 4) extract only the strongest features, whilst larger, unnatural values extract more features and produce a transformed image with less noise. In the exponential model a similar mechanism works, although inversely, so that smaller values of k make rays adhere very closely to features. Common values used for n_{\max} were 4 or 40 whilst for k , values between 5 and 100 were common.

Two parameters, l and d have values related to the scale of the image and the scale of the features required. The maximum depth parameter d controls the number of refractions and reflections that a ray may undergo before it is no longer followed. It is the primary means of ensuring a ray is not too computationally expensive by stopping it after a number of changes of directions. It is usually equivalent to the number of pixels that the

ray has passed through, as in most natural images there is a change in intensity and therefore refractive index between nearly every pixel. It should be set at a value that is large enough to allow a ray to find a structural feature and to follow it, but not so large as to lengthen the time of the transform unnecessarily. Investigation has shown these two properties are best balanced at $d = 256$, although a lower value may be appropriate if the image dimensions are smaller than this value.

The maximum length, l , is primarily present in order to deal with cases in synthetic images where large parts of the image have the same intensity, and hence refractive index. In these cases there may not be many changes in direction and the ray must be terminated after a distance has been travelled. An appropriate value is twice the length of the image diagonal so that it is unlikely to be exceeded before d .

The number of rays that should be traced, N , is an important parameter for ensuring that the image has been sufficiently explored by the rays. It must be set to be large enough for rays to have fully covered the image and reduced the amount of noise, but small enough so as to not generate more rays than are necessary. A complete approach, covering every pixel and a subset of directions is infeasible (requiring at least 100 times as many rays) but it was found that $N = 20000$ was an acceptable heuristic on images with dimensions of 512×512 and below. However, it is more appropriate to use a stopping condition to decide when the transformation of an image is complete, and this is described in section 2.3.2.

Computation time depends upon selected parameters to a certain extent, but for an image of size 512×512 , with standard parameters, computation was recorded to take 1.93s. Whilst this is longer than techniques such as

Sobel, which took 0.022s on the same machine, it is still far from being computationally expensive. In addition to this the technique is easy to parallelise with each ray being independent of the other rays. It should also be noted that the additional methods used in section 3, such as the Hough transform, template matching and hysteresis thresholding all have a computational cost significantly greater than the IRT.

2.3. Refinements

Whilst the image ray transform as described is a powerful technique, two further additions can be made that expand its applications and increase its ease of use.

2.3.1. Intensity and Refractive Index

The IRT, as described in section 2.2, will only emphasise features that are of a higher intensity than their surrounding area. This occurs because total internal reflection only works to stop rays moving from higher to lower refractive indices (hence intensities) and not vice versa. Different measures can be taken to deal with this, depending upon the desired result.

When the approximate intensity of the desired structures are known, then a simple transformation can be applied to make that intensity (the target intensity) the largest value. We do this by finding the difference between the target intensity τ and the original intensity i_o for all pixels in the image:

$$i_\tau = |i_o - \tau| \tag{7}$$

Further to this, if the intensity of the structures is not known, or varies, the IRT can be performed multiple times, with different target intensities

and the results combined by selecting the maximal value at each pixel. A specific and useful case is the selection of values such that $\tau = \{0, 255\}$ (or the original and inverted image) allowing structures that are both high surrounded by low, and low surrounded by high intensity to be highlighted.

2.3.2. Automatic Stopping Condition

The issue of selecting the optimum number of rays that should be followed to ensure the transformed image is of sufficient quality is an important one. An automatic system for stopping the transform when this level has been reached is therefore preferable to using fixed values of N . One method to do this is to calculate the difference in the normalised accumulator between iterations and cease when there is no longer any significant change. A simple and effective method of doing this is the root mean squared (RMS) difference between the intensities of the normalised accumulator at different iterations:

$$D^{<t>}(\mathbf{I}'^{<t>}, T) = \sqrt{\frac{1}{|\mathbf{I}'|} \sum_{z \in \mathbf{I}'} (\mathbf{I}'^{<t>}(z) - \mathbf{I}'^{<t-T>}(z))^2} \quad (8)$$

where T is the number of iterations between each comparison and $\mathbf{I}'^{<t>}$ is the normalised accumulator at iteration t . As the normalisation process is significantly slower than an iteration of the transform, it is appropriate to use large values of T , to balance the computational cost of normalisations with that of performing too many iterations and obtaining a value that is sufficiently robust to noise. The value, D , that is calculated can be employed as a stopping condition by ceasing the transform when it drops below a set limit, D_S . A suitable value of D_S is 1, when $T = 1000$, as this corresponds to a mean intensity change per pixel of 1 per 1000 iterations.

```

while  $D > D_S$  do
    accumulator  $\leftarrow$  CastRandomRay(image, accumulator);
    if  $i \% T = 0$  then
        previousNormImage  $\leftarrow$  currentNormImage;
        currentNormImage  $\leftarrow$  NormaliseArray(accumulator);
         $D \leftarrow 0$ ;
        for  $z \leftarrow 0$  to Size(image) do
             $D \leftarrow D + (\text{currentNormImage}[z] - \text{previousNormImage}[z])^2$ 
        end
         $D \leftarrow \text{SquareRoot}(D / \text{Size}(\text{image}))$ 
    end
     $i \leftarrow i + 1$ ;
end

```

Figure 2: Pseudocode showing the application of the Root Mean Squared difference stopping condition.

Figure 2 shows a pseudocode implementation of the stopping condition. The majority of the IRT occurs in the `CastRandomRay` function and is not detailed, but the calculation of D every T iterations is shown.

2.4. Beams

The image ray transform can be extended to use beams, that is, multiple rays arranged in a line orthogonal to their direction. The addition of this mechanism produces a technique that is able to overcome noise successfully and extract larger features from images. Assuming \mathbf{p}_0 and \mathbf{V} are the starting position and direction of a ray, assigned randomly as in the standard IRT, and $\mathbf{N}_\mathbf{V}$ is the normal to \mathbf{V} then the starting positions of the B rays \mathbf{p}_b where $0 \leq b < B$ that constitute the beam, also of width B , are

$$\mathbf{p}_b = \mathbf{p}_0 + b \cdot \mathbf{N}_\mathbf{V} \quad (9)$$

This alone does not add anything to the IRT. The rays must be followed in parallel and they must be forced to maintain a single beam. This can be done by introducing a tether, a step where the directions of all rays that compose the beam are adjusted towards the beam's mean.

$$\mathbf{V}'_b = (1 - \gamma_v) \mathbf{V}_b + \gamma_v \left(\frac{1}{B} \sum_{q=0}^{B-1} \mathbf{V}_q \right) \quad (10)$$

The parameter γ_v controls the rate at which the direction moves towards the mean, and so controls the extent to which individual rays can diverge from the beam. When $\gamma_v = 0$ no directional adjustment occurs whilst when $\gamma_v = 1$ all directions are set to the mean direction after every iteration. With an intermediate value such as $\gamma_v = 0.5$ the beam is able to change direction if the majority of the rays comprising the beam are changing in

the same direction. An example of the use of beams to transform noisy images is shown in section 3.2.1.

3. Structural Feature Detection

Figure 3 shows the result of the standard ray transform on a variety of images from the Caltech101 dataset (Fei-Fei et al., 2007). Throughout these images we vary the value of target intensity parameter τ in order to highlight the most prominent structural feature clearly. We also employ the stopping condition extension to automatically cast more rays for images of greater size or containing weaker structure.

In figure 3(a) the transform’s propensity for emphasising tubular features is shown clearly. It performs strongly at this task because tubular features act like optical fibres, where a high refractive index glass cylinder surrounded by low refractive index air ensures that rays collect inside them and reflect along their length. In sections 3.2 and 3.3 two applications of tubular structure detection will be discussed. Figure 3(b) show the transform reacting to circular features within images. A ray exiting a circular structure will reflect off the edge at an angle that is shallow enough to ensure that it continues to reflect off the edges around the entire circle. A discussion of further tests on this application is presented in section 3.1.

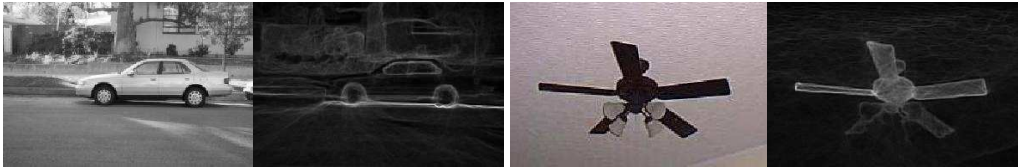
Some of the limitations of the IRT’s application to structural feature detection are shown in figure 3(c). The difficulty of directing the IRT to emphasise only tubular or circular structures is shown in the image of a car. The IRT might be expected to highlight the car wheels most strongly as they are a strong circular feature, but the shadow of the car produces a stronger tubular structure, directing rays away from the wheels. The results



(a) Tubular Features



(b) Circular Features



(c) Examples of Limitations of the IRT

Figure 3: The result of the transform in highlighting structural features in a range of images. $D_S = 1$, $n_{\max} = 40$, $d = 256$

of the transform can also be affected by the other, undesired structures in the image. In the second image of the ceiling fans we aim to highlight the fan blades equally, but the occlusion of two of the blades reduces the response across the whole of those blades, rather than just the occluded area.

3.1. Circle Detection

The standard method for circle detection is the Hough transform for circles, and the image ray transform can be employed to improve results (Cummings et al., 2010a). The Hough transform (HT) (Hough, 1962) was originally used for detecting lines, but can be applied to the detection of arbitrary shapes, especially circles. Yuen et al. (1990) review a number of varieties of the HT for circles, but the faster, decomposed version that was selected for use here was described by Davies (1984).

A standard procedure for performing the HT is to use an edge detection technique, followed by thresholding and then application of the HT on the binary image. Through use of the IRT as a preprocessor before edge detection, we can enhance circular features so that they produce stronger edges and lead to circles being detected with greater accuracy. In order to test this, a number of synthetic images were generated by placing circles of random size and intensity on a variety of natural and synthetic backgrounds (figure 4). The methods described previously (edge detection both with and without the IRT) were then applied to the images and the distance from the ground truth position and radius was found. Two edge detection techniques were used: Sobel and Canny. Through use of the two pass aggregation method described in section 2.3.1 (where $t = 0, 255$) circles of both

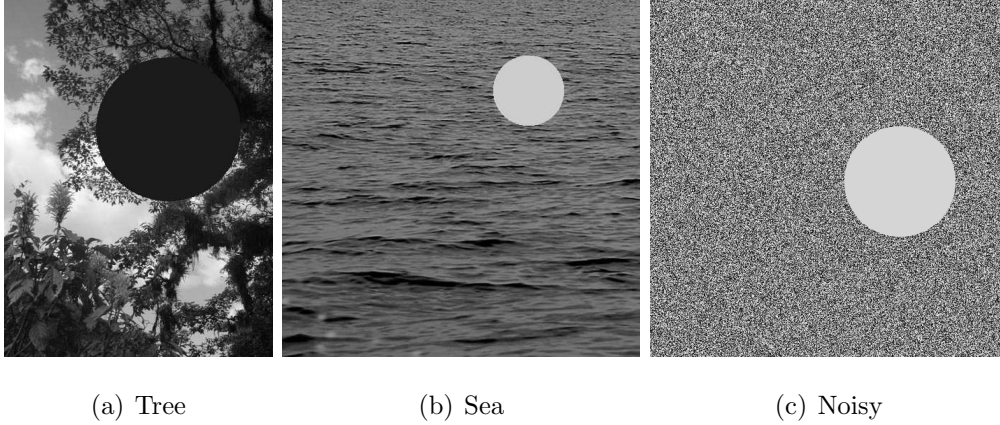


Figure 4: Examples of images generated on different backgrounds.

high and low intensity could be found equally well.

Across all three image types, error was reduced with both types of tested edge detector. The method was also more capable of dealing with situations where the background was of similar average intensity to the circle than the edge detector alone, although performance was weakest in such cases.

3.2. *Enrolment for Ear Biometrics*

Ears are an encouraging area of biometrics and research suggests that they may have similar identification strength to faces (Chang et al., 2003). In addition to this, they exhibit advantages over other biometrics; in particular the ear is relatively immune to the effects of ageing. Whilst identification by ears is obviously an important part of ear biometrics research, of similar importance is the process of extracting and normalising the original image so that accurate identification can occur. This process is known as enrolment and for most identification experiments is currently done manually. There has not been as much research into the automation of detection

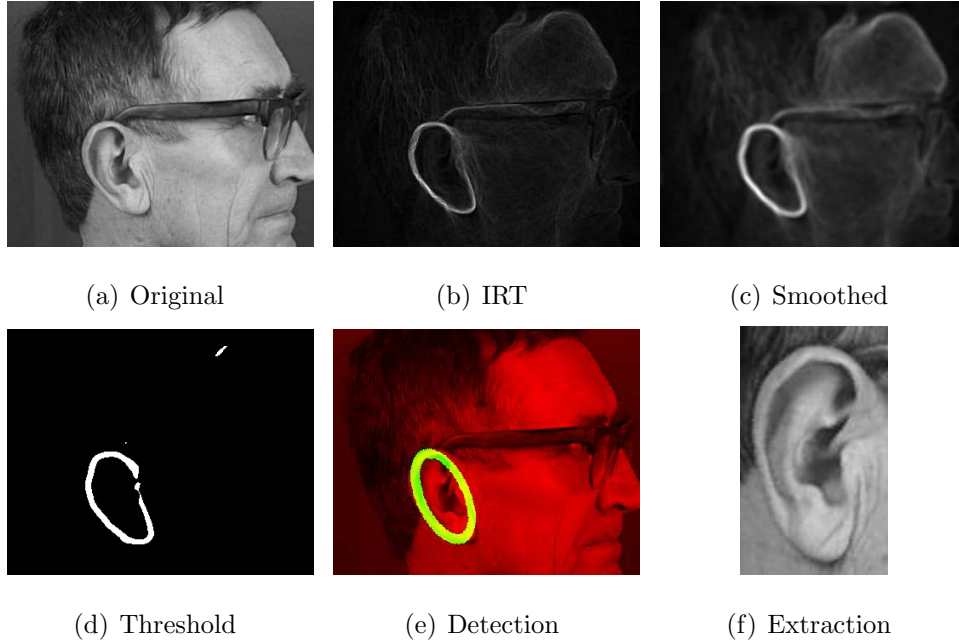


Figure 5: Example of the steps taken to achieve successful ear enrolment.

and enrolment as identification, but attempts have been made using active contours and an ovoid model (Alvarez et al., 2005), an elliptical Hough transform (Arbab-Zavar and Nixon, 2007), SIFT feature points to transform an image to a model (Bustard and Nixon, 2008) as well as Haar wavelets combined with AdaBoost (Islam et al., 2008). The image ray transform can emphasise tubular structures within an image and due to the shape of the ear, this can be exploited to enable enrolment (Cummings et al., 2010b). The outer rim of the ear, known as the helix, is an example of a tubular structure that the IRT is adept at emphasising. The ear is also bounded by the helix, so that by finding it the position of the ear is also found.

The method used to extract the ear image is shown in figure 5. The original image from the XM2VTS database (Messer et al., 1999) undergoes

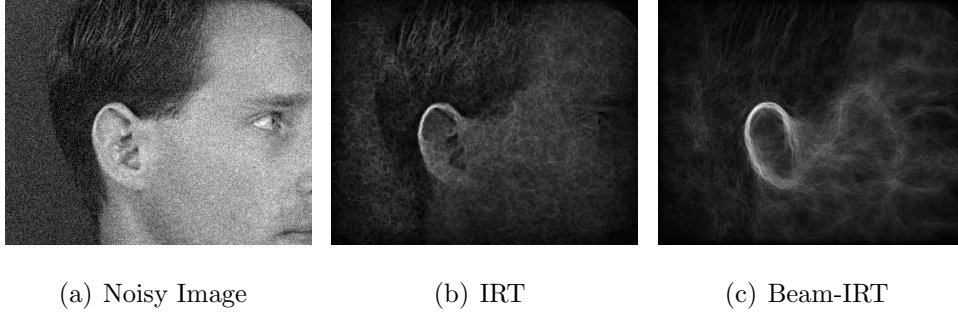
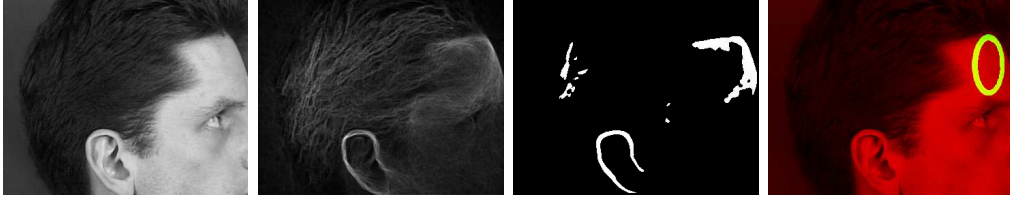


Figure 6: Noisy ear image transformed with the versions of the IRT.

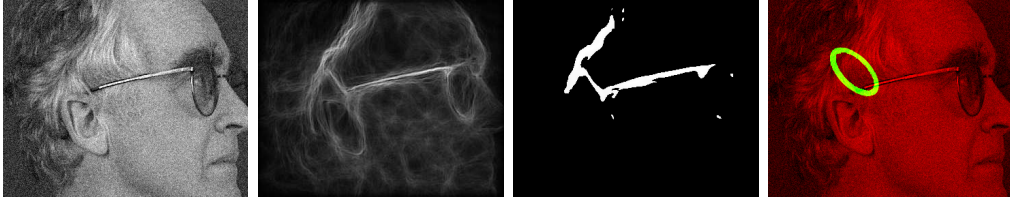
the image ray transform (figure 5(b)). Gaussian smoothing is then applied across the image (figure 5(c)) in order to reduce noise and remove gaps in the helix. The image is then thresholded (figure 5(d)) to remove as much of the image as possible whilst leaving the helix intact before an elliptical template is matched to the image (figure 5(e)). Finally the position, scale and rotation of the ear, found from template matching, are used to extract and normalise the ear (figure 5(f)).

Across 252 images, 99.6% detection was achieved. Of these, 98.4% were correctly normalised for position, scale and rotation. The extension of the transform to beams described in section 2.4 can be used to improve the robustness of the ear enrolment technique to noise. Gaussian noise ($\sigma = 20$) was added to the 252 ear images and the tests repeated as previously described, but using the beam IRT with parameters $B = 8$ and $\gamma_v = 0.5$ (shown in figure 6). The use of the beam transform on noisy images increased the rate of detection of ears from 93.3% to 96.8%, and rate of correct normalisation of detected ears from 64.7% to 93.9%.

The image ray transform and the beam variant both caused the enrolment technique to fail in similar cases. In some cases the transform response



(a) IRT Failure



(b) Beam IRT failure

Figure 7: Examples of failure of the ear enrolment technique.

to other areas of the image was large enough to cause the template to be matched to an incorrect region. Figure 7(a) shows the only example of this detection failure on the original ear images. Whilst the IRT emphasises the ear strongly, it has similarly done so with the forehead, to the extent that it matches more strongly with the template. With the addition of noise both the IRT and its beam variant failed in detection more frequently. In figure 7(b) an example of combination of other structures creating one that responds strongly to both the IRT and the template matching is shown. The spectacle frames and white areas of hair are both highlighted strongly by the transform and the template matching step, to a much greater degree than the ear. There a number improvements to the enrolment technique that could be made to deal with these issues, including removal of straight lines prior to thresholding and the use of a superior matching technique, such as the Hough transform.

3.2.1. Recognition on enroled images

We quantitatively test the ear enrolment technique by using principal components analysis (Chang et al., 2003) (PCA or “eigen-ears”) for recognition on enrolment by the IRT. PCA recognition is a powerful technique but is reliant on accurate ear registration, making it well suited to test the IRT-based enrolment technique. We use a standard PCA implementation (Beveridge et al., 2005) as used previously for ear recognition (Chang et al., 2003; Yan and Bowyer, 2005). Images are normalised for position, scaled to be the same as that of the template (80×150) and are rotated to be vertical using the parameters found during template matchings. During PCA we retain 60% of the eigenvectors.

We compare our recognition results on the 63 subjects of the XM2VTS database (selected by Hurley et al. (2005)) with those of the automated elliptical Hough transform based enrolment system of Arbab-Zavar and Nixon (2007) as well as the results on manually enroled images. Table 1 shows the rank one recognition rate on PCA with the Manhattan distance (except where stated otherwise), using one of the subjects images as a probe and the rest as the gallery. Our method is clearly superior to that of the Hough transform technique and is approaching the quality of manually enroled images. Further improvement can be made through use of the Yambor Angle metric (Yambor et al., 2002).

Recognition was also tested on noisy images using both the IRT and the beam variant. The averaging effect of the beams caused the rank one recognition rate to be increased from 44.4% to 74.6%.

Enrolment Method	Recognition Rate
Manual (Hurley et al., 2005)	98.4%
IRT (Yambor Angle)	95.2%
IRT	93.7%
Hough Transform (Arbab-Zavar et al., 2007)	75.1%

Table 1: Rank one recognition rates of different enrolment techniques using PCA.

3.3. Retinal Vasculature Detection

The automation of the detection of retinal vasculature is a task that can aid diagnosis of a number of diseases, including diabetic retinopathy. Chaudhuri et al. (1989) were the first to approach the problem, using matched filters. Zana and Klein (2001) used a morphological approach to the problem and Jiang and Mojon (2003) examined different ways of thresholding. The DRIVE database was assembled by Staal et al. (2004) whilst they also developed a supervised learning method of pixel-wise classification based upon ridges in the image. Soares et al. (2006) created another supervised technique based upon Morlet wavelets and Niemeijer et al. (2004) tested previous unsupervised techniques upon the DRIVE database. Blood vessels are another example of a tubular structure that the image ray transform can successfully highlight, and with simple techniques vessels can be segmented successfully (Cummings and Nixon, 2010).

The method through which this occurs is shown in figure 8. An image from the DRIVE database is taken (figure 8(a)), and the green channel extracted (figure 8(b)), a standard step in order to use the channel with greatest contrast. The IRT is then applied (figure 8(c)) to this image, and some simple automated processing removes some undesired features from

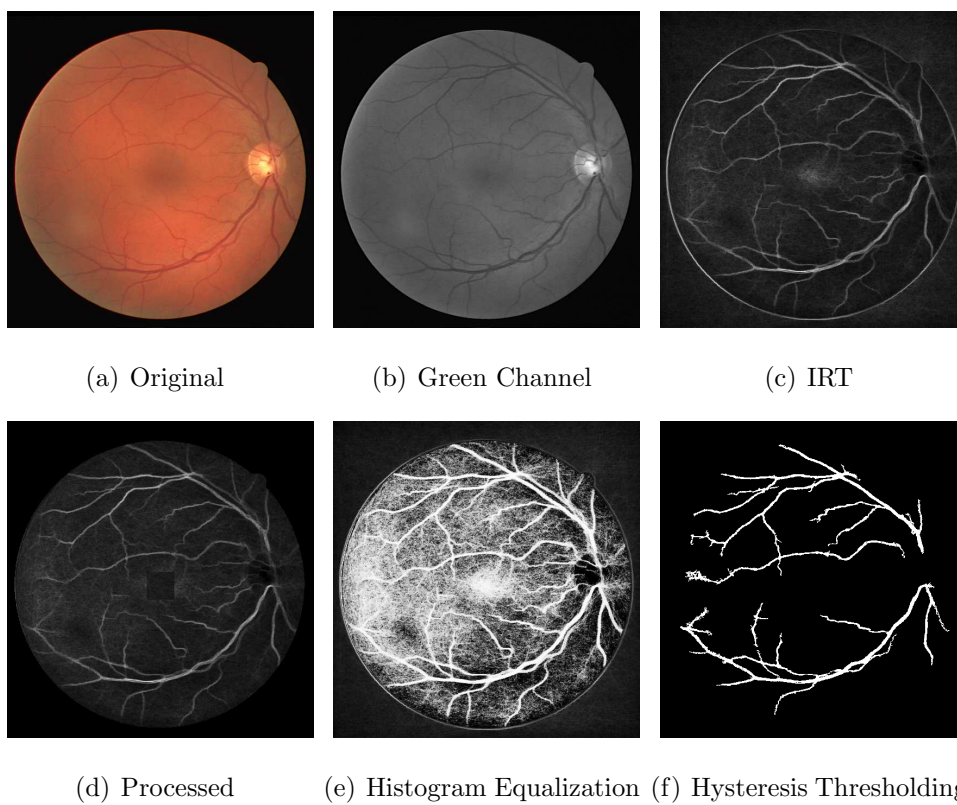


Figure 8: Steps taken to extract retinal blood vessels.

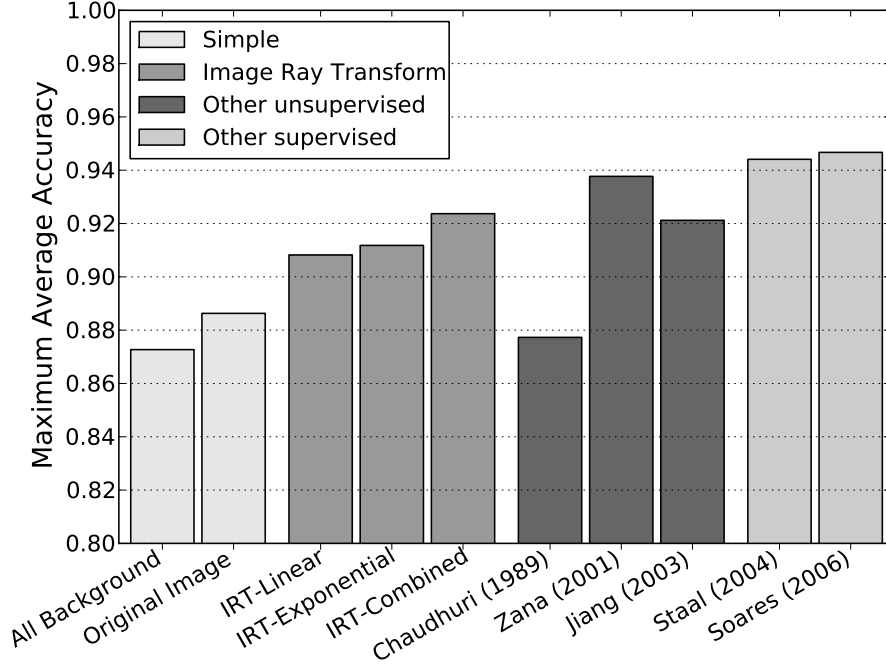


Figure 9: Results of the IRT based technique compared with a selection of other results.

the result (figure 8(d)). This is followed by a simple method of segmenting the image; histogram equalisation (figure 8(e)) and Hysteresis Thresholding (figure 8(f)). This thresholded image is then compared in a pixel-wise fashion with the ground truth from the database.

A number of versions of the transform were tested, but the best results were found from combining the results of the transform with linear refractive indices and the transform with exponential refractive indices, as they complemented each others strengths and weaknesses well. One problem found was that the transform would regularly highlight the fovea (the light central patch in figure 8(c)), and the border of the retina which had to be automatically removed in order to prevent incorrect segmentation. Figure

9 compares the maximum average accuracy (mean percentage of correct pixels across 20 test images) of the IRT technique with others. We compare with the some of the other techniques described previously, as well as those achieved when all pixels are set to the most populous class (all background) and performing histogram equalisation and hysteresis thresholding on the green channel intensity image (original image). Results are better than naïve techniques, comparable with other unsupervised techniques, but not as good as supervised learning techniques, as would be expected.

4. Conclusions and Future Work

The image ray transform is a unique approach to feature extraction that can aid the detection of structural features. It can enhance the performance of circle detection techniques through its natural propensity to emphasis circles. Tubular structures such as the helix of the ear and the blood vessels within the retina can be enhanced so that they can be extracted by even simple techniques. We have shown that through the use of beams as local averaging the IRT can be improved so that it is more robust to noise.

Looking forward, there are many areas in which the image ray transform could be applied and extended. The detection of structural features is an integral part of a wide and varied range of applications. The ear biometric enrolment technique could be enhanced through a better matching algorithm (such as an elliptical Hough transform) or through the development of an identification method based upon the IRT. Another biometric that might be appropriate for application of the transform is that of gait as legs are well emphasised tubular structures. Through application of the IRT as a preprocessor for more advanced segmentation techniques for retinal vessel

extraction, results that are not merely comparable, but superior to other techniques should be possible. In addition, there are numerous other areas of biomedical imaging in which the IRT can aid detection.

- Alvarez, L., Gonzalez, E., Mazorra, L., 2005. Fitting ear contour using an ovoid model, in: 39th Int'l Carnahan Conf. on Security Technology (CCST05).
- Arbab-Zavar, B., Nixon, M., 2007. On shape-mediated enrolment in ear biometrics, in: 3rd Int'l Symp. on Visual Computing (ISVC07).
- Arbab-Zavar, B., Nixon, M., Hurley, D., 2007. On model-based analysis of ear biometrics, in: IEEE Conference on Biometrics: Theory, Applications and Systems.
- Beveridge, J., Bolme, D., Draper, B., Teixeira, M., 2005. The CSU face identification evaluation system. *Machine vision and applications* 16, 128–138.
- Bustard, J.D., Nixon, M., 2008. Robust 2d ear registration and recognition based on sift point matching, in: 2nd IEEE Int'l Conf. on Biometrics Theory, Applications Systems (BTAS 08).
- Chang, K., Bowyer, K., Sarkar, S., Victor, B., 2003. Comparison and combination of ear and face images in appearance-based biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25, 1160–1165.
- Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M., Goldbaum, M., 1989. Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE Transactions on Medical Imaging* 8, 263–269.
- Cummings, A.H., Nixon, M.S., 2010. Retinal vessel extraction with the image ray transform, in: 6th Int'l Symp. on Visual Computing (ISVC10).
- Cummings, A.H., Nixon, M.S., Carter, J.N., 2010a. Circle detection using the image ray transform, in: Int'l Conf. Computer Vision Theory and Applications (VISAPP 2010).
- Cummings, A.H., Nixon, M.S., Carter, J.N., 2010b. A novel ray analogy for enrolment of ear biometrics, in: 4th IEEE Int'l Conf. on Biometrics Theory, Applications Systems (BTAS 10).
- Davies, E.R., 1984. A modified Hough scheme for general circle location. *Pattern Recognition Letters* 7, 37–43.
- Fei-Fei, L., Fergus, R., Perona, P., 2007. Learning generative visual models from few

- training examples: An incremental Bayesian approach tested on 101 object categories. *Computer Vision and Image Understanding* 106, 59–70.
- Hill, F., 2000. *Computer graphics using OpenGL*. Prentice Hall. chapter 12. 3rd edition. p. 678.
- Hough, P., 1962. Method and means for recognizing complex patterns. U.S. Patent 3.069.654.
- Hurley, D.J., Nixon, M.S., Carter, J.N., 2005. Force field feature extraction for ear biometrics. *Computer Vision and Image Understanding* 98, 491–512.
- Islam, S., Bennamoun, M., Davies, R., 2008. Fast and fully automatic ear detection using cascaded adaboost, in: *Proc. of IEEE Workshop on Application of Computer Vision WACV*, pp. 1–6.
- Jiang, X., Mojon, D., 2003. Adaptive local thresholding by verification-based multi-threshold probing with application to vessel detection in retinal images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25, 131 –137.
- Messer, K., Matas, J., Kittler, J., Luettin, J., Maitre, G., 1999. XM2VTSDB: The extended M2VTS database, in: *2nd Int’l Conf. on Audio and Video-based Biometric Person Authentication (AVBPA 1999)*.
- Meyer, F., Maragos, P., 1999. Multiscale morphological segmentations based on watershed, flooding, and eikonal pde. *Lecture Notes in Computer Science* 1682, 351–362.
- Niemeijer, M., Staal, J., van Ginneken, B., Loog, M., Abramoff, M., 2004. Comparative study of retinal vessel segmentation methods on a new publicly available database, in: *Proceedings of SPIE*, p. 648.
- Nixon, M.S., Liu, X.U., Direkoglu, C., Hurley, D.J., 2011. On using physical analogies for feature and shape extraction in computer vision. *The Computer Journal* 54, 11–25.
- Pnueli, Y., Bruckstein, A., 1994. Digidürer - a digital engraving system. *The Visual Computer* 10, 277–292.
- Soares, J., Leandro, J., Cesar Jr, R., Jelinek, H., Cree, M., 2006. Retinal vessel segmentation using the 2-d morlet wavelet and supervised classification. *IEEE Transactions on Medical Imaging* 25, 1214–1222.
- Staal, J., Abramoff, M., Niemeijer, M., Viergever, M., van Ginneken, B., 2004. Ridge based vessel segmentation in color images of the retina. *IEEE Transactions on Medical*

- Imaging 23, 501–509.
- Yambor, W.S., Draper, B.A., Beveridge, J.R., 2002. Analysing PCA-based face recognition algorithms: Eigenvector selection and distance measures, in: Empirical Evaluation Methods in Computer Vision. World Scientific.
- Yan, P., Bowyer, K.W., 2005. Empirical evaluation of advanced ear biometrics. Computer Vision and Pattern Recognition Workshop 0, 41.
- Yuen, H.K., Princen, J., Illingworth, J., Kittler, J., 1990. Comparative study of Hough transform methods for circle finding. Image and Vision Computing 8, 71–77.
- Zana, F., Klein, J., 2001. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. IEEE Transactions on Image Processing 10, 1010–1019.