Colour normalisation of retinal images

Keith A Goatman[a*], A David Whitwam[a], A Manivannan[a], John A Olson[b], and Peter F Sharp[a]

[a]Department of Bio-Medical Physics and Bioengineering, Aberdeen University and
[b]Eye Clinic, Aberdeen Royal Infirmary, Foresterhill, Aberdeen, AB25 2ZD

Abstract. The development of a nationwide eye screening programme for the detection of diabetic retinopathy has generated much interest in automated screening tools. Currently most such systems analyse only intensity information — discarding colour information if it is present. Including colour information in the classification process is not trivial; large natural variations in retinal pigmentation result in colour differences between individuals which tend to mask the more subtle variation between the important lesion types. This study investigated the effectiveness of three colour normalisation algorithms for reducing the background colour variation between subjects. The normalisation methods were tested using a set of colour retinal fundus camera images containing four different lesions which are important in the screening context. Regions of interest were drawn on each image to indicate the different lesion types. The distribution of chromaticity values for each lesion type from each image was plotted, both without normalisation and following application of each of the three normalisation techniques. Histogram specification of the separate colour channels was found to be the most effective normalisation method, increasing the separation between lesion type clusters in chromaticity space and making possible robust use of colour information in the classification process.

1 Introduction

Diabetic retinopathy is currently the major cause of blindness in the UK working-age population. The fact that blindness can usually be delayed and often prevented, providing the disease is caught sufficiently early, has recently prompted the establishment of a nationwide screening programme. Since approximately 2% of the population are diabetic, and annual screening has been recommended, the screening programme will generate a very large number of images for analysis. It is therefore not surprising that interest in automated screening techniques [1–3] has increased rapidly in the last few years. However, despite high resolution colour cameras being the accepted standard for screening programmes, automated software tends currently to base its analysis on intensity information alone, either from 'red-free images' or using the green channel of RGB colour images. More than a decade ago Goldbaum et al. (1990) [4] showed significant differences in the colour measurements of lesions in retinal images. Since then little interest has been shown in colour classification of retinal images. In practice, while models exist to identify abnormal coloured objects within the retinal image [5,6], without some form of colour normalisation or adaptation for the background pigmentation the large variation in natural retinal pigmentation across the population confounds discrimination of the relatively small variations between the different lesion types.

The human visual system is a poor spectral analyser; our perception of colour is based on the responses of only three receptor types sensitive to three bands of wavelengths. The consequence of this is that widely differing spectra produce the exactly the same colour perception. Colour cameras also use only three receptors, since this is all that is required to match human perception of colour. Given the remitted spectrum it is possible to calculate the red, green and blue colour channels values. However the inverse problem is hugely under-determined, hence changes which may be deduced using a multi-channel spectrum analyser will not necessarily be detectable using only three colour sensors. A feature of human vision is that it adapts automatically and subconsciously to relatively large changes in the illuminating spectrum so that white objects are still perceived as being white. A similar effect is seen, for instance, if an image is projected onto a screen which is cream coloured; white objects in the projected image are still perceived as being white. This process is known as *colour constancy*. In contrast to human vision, colour cameras do not adapt automatically to changes in illumination. The lesion colour measured by the camera depends on:

1. **Lesion composition:** All the lesions are composed of different materials with different reflection, absorption, and scattering properties.
2. **Lesion density:** All lesions vary in their size and thickness. The density of the lesion controls how much light is transmitted/reflected by the lesion (i.e. the colour can vary from the pure lesion colour to almost the retinal background colour).

*Email: k.a.goatman@biomed.abdn.ac.uk

3. **Scattered/reflected light:** The colour and intensity of light scattered and/or reflected within the retina itself (probably negligible in a healthy, bleached retina) and the orbit.

4. **Lens colouration:** The lens becomes increasingly yellow (absorbing blue wavelengths) with age above around 30 years.

Note that all the lesions, except drusen, are positioned in front of all the pigmented retinal tissue (i.e. in front of the RPE, choroid and photoreceptors). All the lesions are of a similar colour and occupy a relatively small area of the complete colour space.

2 Method

Three colour normalisation algorithms originally intended for making colour images invariant with respect to the colour of the illumination were investigated for their ability to make the retinal images invariant with respect to background pigmentation variation between individuals. Colour normalisation does not aim to find the true object colour, but to transform the colour so as to be invariant with respect to changes in the illumination — without losing the ability to differentiate between the objects of interest. The three methods tested were:

1. **Greyworld:** The greyworld normalisation assumes that changes in the illuminating spectrum may be modelled by three constant multiplicative factors applied to the red, green, and blue channels. Since the mean values of the red, green, and blue channels will be multiplied by the same constants dividing each colour channel by the respective mean value removes the dependence on the multiplicative constant. An iterative variation of the greyworld normalisation [7] (which includes intensity normalisation) was not found to perform significantly better.

2. **Histogram equalisation:** Histogram equalisation of the individual red, green, and blue channels represents a more powerful normalisation transformation than the greyworld method [8]. It is based on the observation that for each colour channel pixel rank order is maintained under different illuminants, i.e. if under one illuminant the red values of two pixels are r_1 and r_2, where $r_1 < r_2$, then under another illuminant, although the magnitudes of r_1 and r_2 may change, r_1 should still be less than r_2 (there are, however, some conditions where this will not be true). Histogram equalisation is a non-linear transform which maintains pixel rank and is capable of normalising for any monotonically increasing colour transform function. The proportion of the different tissue types must be similar in all images to be normalised. Equalisation tends to exaggerate the contribution of the blue channel (the normal retina reflects little blue light).

3. **Histogram specification:** Histogram specification [9] transforms the red, green, and blue histograms to match the shapes of three specific histograms, rather than simply equalising them. This has the advantage of producing more realistic looking images than those generated by equalisation, and it does not exaggerate the contribution of the blue channel. For this study the reference histograms were taken from an arbitrary normal image with good contrast and coloration. Histogram specification has been used before for normalising retinal colour to aid the detection of hard exudates [10].

In order to compare the normalisation methods a dataset of 18 colour retinal fundus camera images was compiled, where each image was known to contain at least one of the following lesion types which are important for retinopathy screening:

- **Cotton wool spots (CWS):** Swelling of the nerve fibre layer axoplasm in response to retinal ischaemia, transforming it from transparent to highly reflective, appearing bright (slightly blue) white. They can be very dense (for instance they may completely block fluorescence emanating from beneath them in an angiogram). They have ill-defined edges (hence their name). They usually disappear spontaneously after around 8 weeks.

- **Hard exudates (HE):** Lipid deposits in the inner nuclear layer as a result of vascular leakage. They are highly reflective and appear bright yellow, often with a distinctive spatial distribution.

- **Blot haemorrhages (BH):** Leakage of blood in the inner nuclear layer. They appear dark red.

- **Drusen:** Debris deposited below the retinal epithelium layer (RPE) and collected in Bruch's membrane due to the turnover of retinal receptor pigments. They appear yellow. Although not related to diabetic retinopathy (they are more commonly associated with age related macular degeneration) they have a similar appearance to HE and are therefore a confounding factor in the identification of HE.

The images were acquired using a Topcon fundus camera and recorded on 35 mm colour slide film. The images were digitised (approx. 1000 dpi) using a Nikon Coolscan 4000ED slide scanner, producing RGB colour images with 8 bits per colour channel. The retinal images are circular; masks were generated automatically by simple thresholding of the green colour channel followed by 5×5 median filtering to exclude the dark background from the colour normalisation calculations. Regions of interest were drawn around the different lesions for all the images and masks produced with a specific greylevel value representing each lesion type. Five of the images contained CWS, fourteen contained HE, and six contained BH. Only two of the images contained drusen. The same region of interest masks were used to analyse the images before and after normalisation.

Colour may be represented independently of its intensity by dividing the red, green and blue channel values by the sum of the three channels, i.e.

$$r = R/(R+G+B), \qquad g = G/(R+G+B), \qquad b = B/(R+G+B)$$

This reduces the three-dimensional RGB colour space cube to a two-dimensional triangular space (since the third ordinate is always one minus the sum of the other two). The resulting intensity normalised coordinates are known as *chromaticity coordinates*. For each image the average chromaticity coordinate for each lesion type present was calculated. In the chromaticity space, a line between any two points passes through all the colours which may be formed by mixing the colours represented by the end points. In this application the lesion colour may vary from pure lesion almost to the background colour so the different lesion types are expected to radiate from the region of the chromaticity space which represents the background colouration.

3 Results

Figure 1(a) plots the average lesion colours in each image without any normalisation. The ellipses shown are centred on the mean position for each lesion type, with the major axis aligned with the direction of maximum variance (found using the Hotelling transform). The radius of the major axis represents two standard deviations in the direction of that axis. The minor axis length represents two standard deviations in the orthogonal direction. All four lesion chromaticity values are seen to overlap. Figure 1(b) shows the effect of the greyworld normalisation, which partially separates the lesion clusters, in particular differentiating the haemorrhages. Figure 1(c) shows the result of equalisation, which also differentiates the haemorrhages, but appears to increase the overlap in the other lesion types. Finally the result following histogram specification is shown in figure 1(d), which shows the clearest separation of the lesion clusters.

4 Conclusions

Three normalisation techniques were tested on a set of retinal images. Histogram specification was found to be the most effective normalisation method, improving the clustering of the different lesion types, removing at least some of the variation due to retinal pigmentation differences between individuals. Colour classification is not intended to replace existing intensity-based classification but to augment it and improve overall classification accuracy.

It was not anticipated that histogram specification should perform so much better than equalisation. One possible explanation is the exaggerated contribution of the blue component following equalisation, which possibly loses subtle but important differences in the blue values due to equalisation quantisation.

An important question is whether the differences in retinal background pigmentation are modelled acceptably as a variation in the colour of the illumination. While this is a safe assumption for changes due to lens colouration (since all the incident and remitted light are so filtered), it is less so for background pigmentation changes since not all retinal tissues are equally affected (i.e. the only contribution for non-pigmented tissues such as the optic disc and highly reflective lesions is from scattered and reflected light from pigmented tissue). Clearly the model is inadequate for dealing with local pigmentation variations across an individual retina. However, despite these reservations the results appear to show that an average correction is much better than applying no correction to the images.

Variation in colour due to scattering in surrounding tissue and reflections within the orbit can be greatly reduced by imaging using a confocal scanning laser ophthalmoscope (SLO) rather than a fundus camera. Early results using our colour SLO [11] appear to show much less variation in lesion chromaticity, resulting in less overlap between lesions even prior to normalisation.

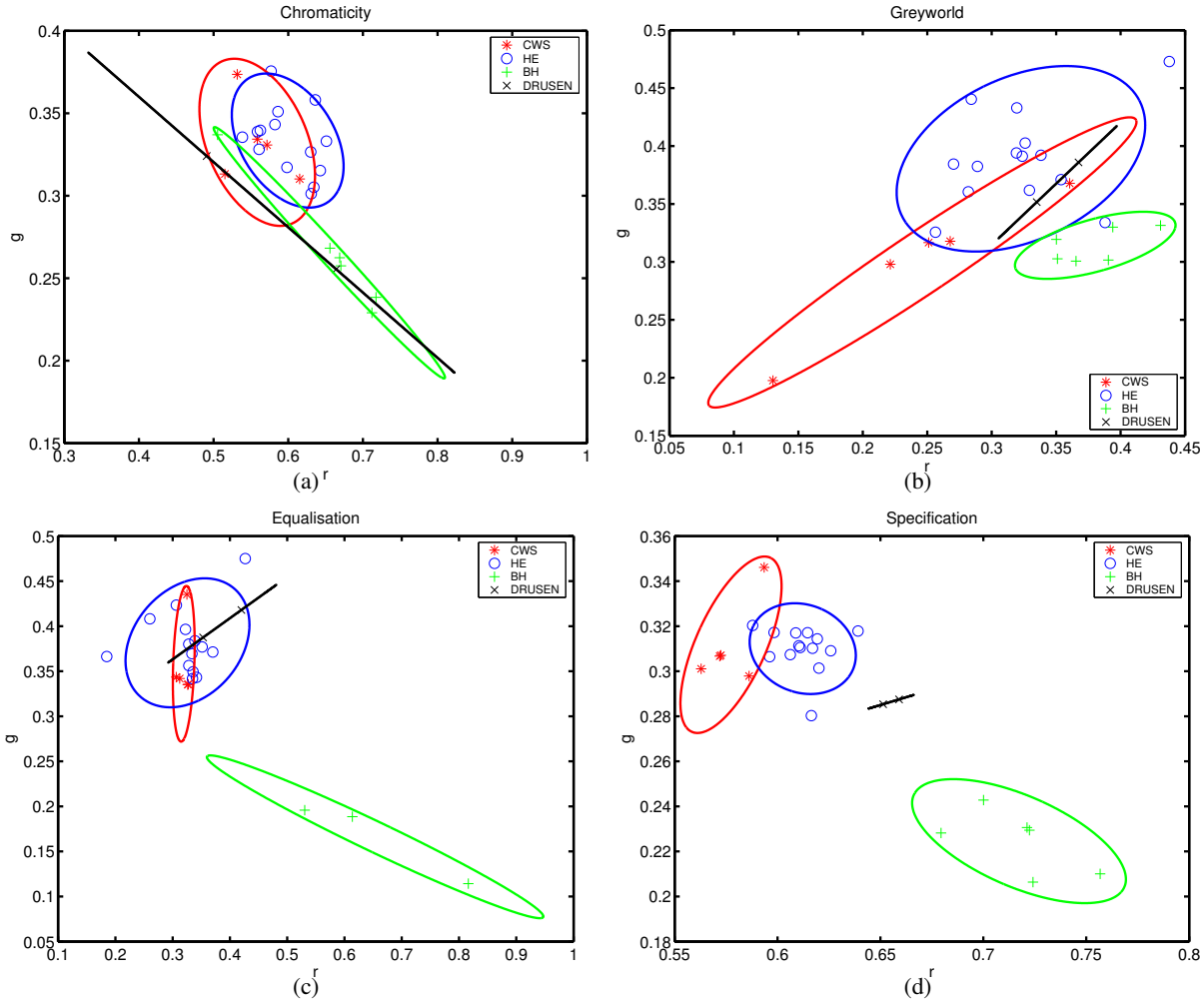

Figure 1. Chromaticity plots: (a) No normalisation, (b) Greyworld normalisation, (c) Histogram equalisation, (d) Histogram specification.

References

1. J. H. Hipwell, F. Strachan, J. A. Olson et al. "Automated detection of microaneurysms in digital red-free photographs: a diabetic retinopathy screening tool." *Diabetic Medicine* **17**, pp. 588–594, 2000.
2. B. M. Ege, O. K. Hejlesen, O. V. Larsen et al. "Screening for diabetic retinopathy using computer based image analysis and statistical classification." *Computer Methods & Programs in Biomedicine* **62**, pp. 165–175, 2000.
3. C. Sinthanayothin, J. F. Boyce, T. H. Williamson et al. "Automated detection of diabetic retinopathy on digital fundus images." *Diabetic Medicine* **19**, pp. 105–112, 2002.
4. M. H. Goldbaum, N. P. Katz, M. R. Nelson et al. "The discrimination of similarly colored objects in computer images of the ocular fundus." *Investigative Ophthalmology & Visual Science* **31**, pp. 617–623, 1990.
5. M. Hammer & D. Schweitzer. "Quantitative reflection spectroscopy at the human ocular fundus." *Physics in Medicine and Biology* **47**, pp. 179–191, 2002.
6. S. J. Preece & E. Claridge. "Monte Carlo modelling of the spectral reflectance of the human eye." *Physics in Medicine and Biology* **47**, pp. 2863–2877, 2002.
7. G. D. Finlayson, B. Schiele & J. Crowley. "Comprehensive colour image normalisation." In *Proceedings of the European Conference on Computer Vision*, pp. 475–490. Springer-Verlag, 1998.
8. S. H. Hordley, G. D. Finlayson, G. Schaefer et al. "Illuminant and device invariant colour using histogram equalisation." Technical Report SYS-C02-16, UEA, 2002.
9. R. C. Gonzalez & R. E. Woods. *Digital image processing*. Prentice Hall, second edition, 2002.
10. A. Osareh, M. Mirmehdi, B. Thomas et al. "Classification and localisation of diabetic-related eye disease." In *Proceedings of the European Conference on Computer Vision*, pp. 502–516. Springer-Verlag, 2002.
11. P. Vieira, A. Manivannan, P. F. Sharp et al. "True colour imaging of the fundus using a scanning laser ophthalmoscope." *Physiological Measurement* **23**, pp. 1–10, 2002.