A Histogram Approach for the Screening of Age-Related Macular Degeneration

Mohd Hanafi Ahmad Hijazi[a*], Frans Coenen[a] and Yalin Zheng[b]

[a] Department of Computer Science, University of Liverpool, Liverpool, UK,
[b] Ophthalmology Research Unit, School of Clinical Sciences, University of Liverpool, Liverpool, UK.

Abstract. The identification of retinal drusen is important in the diagnosis of Age-Related Macular Degeneration (AMD). This is normally undertaken through visual inspection of retinal colour images; a time consuming, resource intensive, process. In this paper an automated approach is proposed to support AMD screening. The fundamental idea is that instead of detecting the physical existence of drusen in the retina, representative patterns of retinal images (with or without drusen) are extracted in the form of histograms. Labelled exemplar histograms are then stored in a "case based". New, "unseen" examples are then classified by comparison with this case base and analysed for drusen using a Dynamic Time Warping (DTW) comparison process. Evaluation using the proposed approach has produced results that are both interesting and promising.

1 Introduction

Age-Related Macular Degeneration (AMD) is the main cause of the elderly blindness in developed countries. A study [1], undertaken in the UK, demonstrated that, between June 1987 and April 2002, approximately 17% of the participants were diagnosed with AMD; further, more than 95% of these were aged 60 years and above. AMD is expected to increase over the coming years. Drusen, yellowish-white sub-retinal deposits located between the retinal pigment epithelium (RPE) of the eye and Bruch's membrane, has been regarded as hallmark of AMD. Identifying and quantifying drusen, through examining of patients' retinal images is essential in diagnosing and staging AMD. For instance, the detection of drusen at the very early stage of AMD is critical for effective treatment options but is a challenging task due to the variety of drusen in size and shape. A substantial amount of research has been undertaken to identify the emergence of drusen through image processing and analysis [2, 3, 4], as well as analysis of alphanumeric medical data [1, 5], which has produced good results. An image registration method for aligning pairs of retinal images was reported in [6], however, expert intervention is still required.

Histograms have been widely used to represent colour distributions in images. They are considered to be a simple way of representing the characteristics of an image in terms of colour distribution within images, and an effective representation for identifying objects in images [9]. Much research has been conducted on the use of histogram as a medium for image retrieval [7, 8]. Generally, there are two methods of generating histograms: (i) *fixed binning* and (ii) *adaptive binning*. Fixed binning applies the same numbers and characters of bins[1] to all images [8], while adaptive binning adapts to the actual distribution of colours in images [9, 10]. Adaptive binning histograms are considered to represent images more efficiently [11], but are constrained by the number of dissimilarity measures available to measure the similarity between histograms. Fixed binning on the other hand can adopt a wider range of similarity metrics (such as Euclidean distance). Each bin in fixed binning histograms can be regarded as a vector and this makes it possible to apply various machine learning and data mining algorithms to the representation for (say) clustering and classification.

In this paper a fixed binning histogram based approach for the automated screening and diagnosis of AMD is proposed. The idea is that instead of detecting the physical existence of drusen on the retina, an approach to identify and extract representative patterns of retinal images (with or without drusen) using histograms is proposed. The main contribution of the approach is that it may serve as a good automated first pass in the screening of retinal images. Dynamic Time Warping (DTW), a technique for mapping pairs of time series curves, is used to measure the similarities between histograms (and consequently retinal images).

2 Backgrounds

The diagnosis of AMD is typically undertaken through the inspection of the macula (see Figure 1). AMD is classified as being either neovascular (wet) or non-neovascular (dry). Neovascular AMD is less common but is more severe than the non-neovascular. The majority of AMD patients who suffer vision loss have the neovascular form of the disease. The presence of drusen is expected with advancing age, people of 40 years of age and above can expect to have some small drusen. However, the presence of larger and more numerous drusen are recognised as an early

[1] colour space cell

E-mail: M.Ahmad-Hijazi@liverpool.ac.uk, coenen@liverpool.ac.uk and yalin.zheng@liverpool.ac.uk

sign of AMD. Drusen are often categorised into two types: (i) *hard* and (ii) *soft* drusen. Hard drusen have a well defined border, while soft drusen have boundaries that often blend into the background. Soft drusen are typically associated with AMD. The advance of retinal image acquiring technology, and the establishment of digital fundus photography, has triggered research into more accurate techniques in measuring and identifying macular drusen [2].

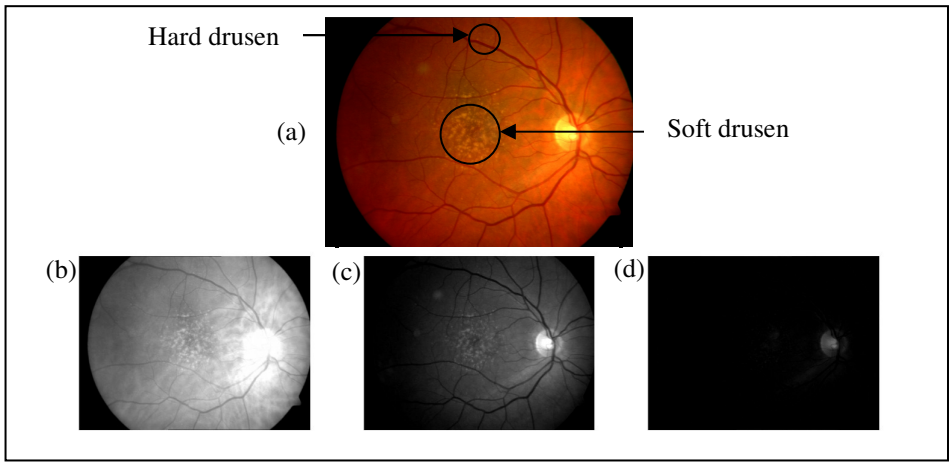

Figure 1. (a) Colour retinal image with drusen circled. (b) red channel image, (c) green channel image, and (d) blue channel image

Image processing techniques have been widely applied to AMD since the 1990's. Much of this work was directed at retina image analysis to improve the effectiveness of drusen detection. Drusen detection represents a significant information technology challenge, and is hampered by a number of difficulties: (i) illumination that causes non-uniformity in the captured images, (ii) object recognition within the images, and (iii) the alignment of images during "capture" (images do not cover identical parts of the retina).

Notable work on automated detection includes the following. In [2] a background levelling technique was applied to fundus retinal images and Focally Increased Auto-Fluorescence (FIAF) to reconstruct the macular background and then remove the background variability from the entire image before the segmentation of the images to identify drusen. In [12] a semi-automated algorithm for drusen detection and segmentation in retinal optical coherence tomography images is described. However, the algorithm in [12] required users to specify the Region of Interest (ROI) of the retina image that is to be segmented. The algorithm has successfully been used to detect drusen within a short time period, but in certain cases failed to detect small drusen that did not change the curve of the RPE (in some cases "false alarms" were generated by mistakenly identifying drusen).

Work on detecting drusen using a histogram-based algorithm has been reported in [3]. The algorithm first enhances a given image by using multilevel histogram equalization (MLE), a modified version of the adaptive histogram equalization algorithm. Once the image is enhanced, the drusen segmentation takes place. Two thresholds are applied, a *global* and a *local* threshold. More specifically Otsu's [13] global thresholding technique was used to carry out the thresholding, while a Histogram-based Adaptive Local Threshold (HALT) was used for the local threshold. Experiments on a very small dataset of 23 images produced good results. Note also that only the green channel was used to represent the images because it is the least effected by illumination. Other work on using histograms for drusen detection is reported in [4], where colour histograms are used to enhance the contrast of drusen against normal retinal pigment epithelium. Their work [4] would be very helpful during the pre-processing stage of the images.

The effectiveness of general histogram based image categorization and retrieval has been empirically measured [7, 8]. One example of the use of histogram for classifying general images can be found in [8] where fixed binning histograms, called a *quasi-histogram*, are used. In [8] quasi-histograms are generated for different image regions and as a state-sequence, rather than a vector. The quasi-histograms are then used in a Markov Chain (MC) process for image classification. Good results are produced in comparison to region based histogram technique used with Hidden Markov Models (HMM) and Support Vector Machines (SVM).

2.1 Dynamic Time Warping

Dynamic time warping (DTW) is a technique for measuring the similarity between two time series sequences. It has been most commonly used in time series analysis [14, 15], but can also be applied in other domains. Thus a histogram, of the form described above, can be interpreted as a time series. DTW uses a dynamic programming

approach to align two time series and then generates a *warping path* that maps (aligns) the two sequences onto each other. To map two time series *T* and *S*, of length *n* and *m* respectively, where $T = t_1, t_2, \ldots, t_n$ and $S = s_1, s_2, \ldots, s_m$ a *n*-by-*m* matrix will be formed, where the (i^{th}, j^{th}) grid point corresponds to the alignment or distance between two points t_i and s_j. The warping path, *W*, is then the set of matrix elements that defines a mapping between *T* and *S*, defined as $W = w_1, w_2, \ldots, w_K$, where $\max(m, n) \leq K < m + n - 1$. The distance $d(t_i, s_j)$ between two points t_i and s_j is used to identify potential warping paths. There are many distance measures that may be used, the most common one is the Euclidean distance, and this is the measure used in this paper. Thus, $d(t_i, s_j) = w_k = (t_i - s_j)^2$. The minimal warping path is selected by calculating the minimum cumulated distance between *T* and *S* as $DTW(T, S) = min\left[\sqrt{\sum_{k=1}^{K} w_k}\right]$.

3 Methodology

The approach to automatically screen retinal images advocated in this paper comprises three phases: (i) image pre-processing, (ii) histogram generation, and (iii) classification. The image pre-processing is required to filter out unnecessary information presented in the retinal images. Thus, the black area that surrounds the retinal images (see Figure 1) is removed. Give *I* images, with rows *R* and columns *C*, *R* and *C* are fixed for each $i \in I$. The pixels information of each image is stored in a matrix *X*. Each element of *X* is referred as $x_{r,c}(i) = \gamma$, where $0 \leq r < R$ and $0 \leq c < C$. γ is the colour value in hexadecimal format for pixel $x_{r,c}$ ($x \in X$) of image $i \in I$.

During histogram generation stage the given retinal images are translated into a histogram representation, *H*. Experiments with two "suites" of histograms are reported in this paper. The first histogram suite comprises histograms (per image) for each of the three colour channels: Red, Green and Blue (RGB). The second suite comprises histograms describing the Hue, Saturation and Intensity (HSI) components for each given retina image. In this work (unlike for example the work described in [3]) all histograms are assumed to carry relevant retinal image information. However it is note that the green and red channels give a better visual contrast of drusen as compared to the blue channel, as shown in Figure 1. The experiment described below, however, demonstrates that some interesting results are also produced by the blue channel histogram. The length of each histogram is fixed to *M* bins, and represented as $h_{i,b}(m) = \beta$, where $h \in H$, $i \in I$, *b* represents the RGB channel or HSI component, $0 \leq m < M$, and β is the histogram value of *m*, normalized to the maximum value recorded by a *m* for a particular $h_{i,b}$. *M* is set to 256 (number of colour space cells) for each red, green and blue histogram, 360 (0^o to 359^o) for hue histogram and 101 (0 to 100) for both saturation and intensity histograms.

Using this histogram based approach a case base of pre-labelled (AMD positive or negative) histograms, $P \subseteq H$, where $P = p_1, p_2, .., p_a$ and *a* is a set of retinal images that have been hand classified by domain experts. The histograms for a new retinal image to be classified, *N*, where $N \in H$and $N \notin P$ are then plotted onto graphs to attain the curves of the histograms, as depicted in Figure 2, before DTW find $p_a \in P$, a histogram that has the best warping path with *N*, of each colour channel or HSI component. Once identified, *N* will be classified into the same class as p_a is.

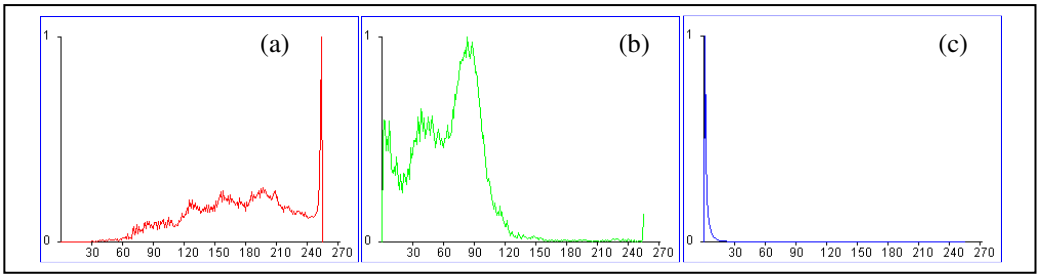

Figure 2. Histogram curves of image in Figure 1, (a) Red channel, (b) Green channel, and (c) Blue channel

4 Experimental Setup

To date the research team have gathered a total of 144, hand labelled images; of which 86 are AMD images and the rest are normal control images collected by the ARIA project[2]. The images are separated into ten equally distributed datasets, with approximately 9 AMD images and 6 control images for each set. Ten-fold Cross Validation was used to evaluate the performance of the proposed approach whereby the image set was divided into 10 subsets and ten evaluations runs conducted. For each run the case base was generated on a different nine tenths and the classification

[2] www.eyecharity.com/aria_online/

accuracy tested on the remaining tenth. The aims of the experimental set up were two-folds: (1) to analyse the effectiveness of using the colour channel histogram, and (2) to investigate the effectiveness of the HSI histograms.

Three evaluation metrics were utilized to measure the classification performance: sensitivity, specificity and accuracy. Sensitivity, tends to measure the effectiveness of the classifier in identifying true positives (AMD images), and is formulated as $sensitivity = \frac{TP}{\alpha}$, where TP (True-Positive) is the number of AMD images classified as AMD by the classifier, and α is the total number of AMD images in the test set. Specificity tries to measure the effectiveness of the classifier in distinguishing the normal control images by not falsely classifying the control image as AMD images. Specificity is defined as $specificity = \frac{TN}{\delta}$, where TN is number of control images not misclassified as AMD images (True-Negative) and δ is the total number of control images in the test set. Accuracy will be used to measure the overall performance of the classifier in term of classifying retinal images correctly according to their class. Accuracy is defined as $accuracy = \frac{TP+TN}{\varepsilon}$, where $\varepsilon = \alpha + \delta$.

5 Results and Discussions

5.1 Performances of Red, Green and Blue Channels and Hue, Saturation and Intensity Components

The results of the experiments using the RGB histogram representation are given in Table 1. The overall performance is promising with a best specificity of 62% (on the green channel) and sensitivity 83% (on the blue channel). Unexpectedly, the highest accuracy of 69% is recorded for the blue channel. Some of the dataset did produce a high effectiveness in both metrics, in particular dataset 1 (red channel).

The results of the experiments using the HSI histogram representation are also given in Table 1. Inspection of HSI columns demonstrates that the results given by the RGB channels are replicated, however with more deficiency in both overall specificity and sensitivity. The best overall performance for specificity, sensitivity and accuracy is achieved via the saturation component, with 60%, 82% and 74% each. Dataset 9 scores the best performance in the hue component.

Data	**Specificity (%)**						**Sensitivity (%)**						**Accuracy (%)**					
	R	**G**	**B**	**H**	**S**	**I**	**R**	**G**	**B**	**H**	**S**	**I**	**R**	**G**	**B**	**H**	**S**	**I**
1	100	80	40	60	60	100	100	56	100	67	100	89	100	64	79	64	86	93
2	33	100	50	83	83	50	88	56	100	75	88	88	64	79	79	79	86	71
3	50	67	67	33	50	50	67	56	100	100	89	67	60	60	87	73	73	60
4	33	50	33	50	67	17	67	44	67	78	56	67	53	47	53	67	60	47
5	50	50	50	50	50	83	78	78	67	56	67	56	67	67	60	53	60	67
6	33	83	50	50	83	33	78	78	100	100	100	78	60	80	80	80	93	60
7	67	83	33	100	67	67	88	63	75	63	88	88	79	71	57	79	79	79
8	20	40	40	20	60	20	78	80	78	89	78	78	57	71	64	64	71	57
9	33	50	67	100	50	33	75	63	75	100	75	75	57	57	71	100	64	57
10	33	50	33	67	17	17	63	50	88	63	100	50	50	50	64	64	64	36
Mean	**47**	***62****	**46**	**57**	**60**	**49**	**76**	**65**	***83****	**81**	**82**	**76**	**65**	**65**	**69**	**72**	***74****	**63**

Table 1. Results of using red (R), green (G), blue (B), hue (H), saturation (S) and intensity (I) for classification of retinal images

5.2 Discussions

With respect to the reported results in Table 1, the most surprising outcome was the performance of the blue channel histograms as compared to the other channels as indicated by the sensitivity measurement. Recall that high sensitivity demonstrates the effectiveness of the classifier in identifying AMD images. This result indicates the important role of the blue channel in classifying retinal images, even when it appears to feature the worst contrast when inspected visually. The classifier has difficulty in the identification of the control images, as shown by the specificity measures. This is due to the fact that there are no consistent patterns of the curves that can be used to identify each class. Further image pre-processing techniques might be required to better prepare the data before any classification process takes place.

The HSI representation results described in Table 2 give similar results to those in Table 1. It is conjectured that the distribution of the histograms will adversely affect performance. Further observation shows that most of the RGB histograms feature an even distribution compared to the HSI histograms. It is conjectured that the uneven distribution

causes the DTW process to calculate an almost similar distance between points in unseen and knowledge base histograms, and consequently inaccurately classified the histogram. However, HSI did better in terms of classification accuracy, with a 74% best accuracy recorded by the saturation component. This shows the ability of HSI in identifying patterns through the colours of the images.

As a comparison, specificity/ sensitivity of 0.81/ 0.70 [2] and 0.99/ 0.98 [3] has been reported in other works on different set of images. It is worth noted however that with refinements on the images visualisation and presentation, a better classifier will be produced, as proved in the previous works [2, 3]. The results reported in this paper are deemed necessary to build up the understanding on the effect of low level image representation that may contribute to the development of a more reliable classifier.

6 Conclusions

In this paper an AMD classifier, founded on a histogram based representation combined with a DTW technique to screen AMD, is proposed. Two types of histograms, RGB and HSI, to represent retinal images were employed to analyse unseen retinal images and to classify these images. The initial results show the superiority of the RGB channels based histograms compared to HSI based histograms.

For future works, the research team intend to identify and learn, using machine learning techniques, more distinct features from the histograms. More advanced image pre-processing techniques, such as image enhancement, segmentation and registration are also to be investigated. With such actions to be taken, it is expected that only the relevant part of an image be represented by histogram and thus the pattern of each class be more consistent. The applicability of other data mining techniques, for instance the association rule mining to mine the statistical information of each image such as its histograms' peak and mean to assist with the classifier learning stage, may also provide a fruitful direction for future work.

References

1. D. Nitsch, I. Douglas, L. Smeeth, , and A. Fletcher "Age-Related macular Degeneration and Complement Activation-related Diseases", *Journal of Ophthalmology* **115**(11), pp 1904-1910, 2008.
2. R.T. Smith, , and U.F. Ahmad "The Role of Drusen in Macular Degeneration and New Methods of Quantification", *Retinal Degenerations: Biology, Diagnostics and Therapeutics*, Eds. J.T. Tink, and C.J. Barnstable, pp 197-211, Humana Press, 2007.
3. K. Rapantzikos, M. Zervakis, and K. Balas "Detection and Segmentation of Drusen Deposits on Human Retina: Potential in the Diagnosis of Age-Related Macular Degeneration", *Journal of Medical Image Analysis* **7**(1), pp 95-108, 2003.
4. L.D. Hubbard, R. P. Danis, M.W. Neider, D.W. Thayer, H.D. Wabers, J.K. White, A.J. Pugliese, and M.F. Pugliese "Brightness, Contrast, and Color Balance of Digital versus Film Retinal Images in the Age-Related Eye Disease Study 2", *Journal of Investigative Ophthalmology and Visual Science* **49**, pp 3269-3282, 2008.
5. V. Dinu, H. Zhao, and P.L. Miller "Integrating Domain Knowledge with Statistical and Data Mining Methods for High-Density Genomic SNP Disease Association Analysis", *Journal of Biomedical Informatics* **40**, pp 750-760, 2007.
6. S.V. Charles "Computer Vision Algorithms for Retinal Image Analysis: Current Results and Future directions", *Lecture Notes in Computer Science: Computer Vision for Biomedical Image Applications* **3765/2005**, pp 31-50, 2005.
7. K.L. Wee, and R. Li "The Analysis and Applications of Adaptive-Binning Color Histograms", *Computer Vision and Image Understanding* **94**, pp. 67-91, 2003.
8. F. Li, Q. Dai, W. Xu, and G. Er "Histogram Mining Based on Markov Chain and Its Application to Image Categorization", *Signal Processing: Image Communication* **22**, pp 785-796, 2007.
9. M.J. Swain, and D.H. Ballard "Color Indexing", *International Journal of Computer Vision* **7**(1), pp 11-32, 1991.
10. X. Gao, B. Xiao, D. Tao, and X. Li "Image Categorization: Graph Edit Distance + Edge Direction Histogram", *Pattern Recognition* **41**, pp 3179-3191, 2008.
11. Y. Rubner, J. Puzicha, C. Tomasi, and J.M. Buhmann "Empirical Evaluation of Dissimilarity Measures for Color and Texture", *Computer Vision and Image Understanding* **84**, pp 25-43, 2001.
12. S. Farsiu, S.J. Chiu, J.A. Izzat, and C.A. Toth "Fast Detection and Segmentation of Drusen in Retinal Optical Coherence Tomography Images", In *Proceedings of Ophthalmic Technologies XVIII*, pp 68440D-1-68440D-12, SPIE, 2008.
13. N. Otsu, "A threshold selection method from gray level histograms", *IEEE Transaction on Systems, Man and Cybernetics* **9**(1), pp 62-66, 1979.
14. D.J. Berndt, and J. Clifford "Using Dynamic Time Warping to Find Patterns in Time Series", *AAAI Workshop on Knowledge Discovery in Databases*, pp 359-370, 1994.
15. E.J. Keogh, and M.J. Pazzani "Derivative Dynamic Time Warping", In *First SIAM International Conference on Data Mining*, 2001.