

Automatic Detection of Blue-White Veil by Discrete Colour Matching in Dermoscopy Images

Ali Madooei, Mark S. Drew, Maryam Sadeghi, and M. Stella Atkins

School of Computing Science
Simon Fraser University
amadooei@cs.sfu.ca,
<http://www.cs.sfu.ca/~amadooei>

Abstract. Skin lesions are often comprised of various colours. The presence of multiple colours with an irregular distribution can signal malignancy. Among common colours under dermoscopy, blue-grey (blue-white veil) is a strong indicator of malignant melanoma. Since it is not always easy to visually identify and recognize this feature, a computerised automatic colour analysis method can provide the clinician with an objective second opinion. In this paper, we put forward an innovative method, through colour analysis and computer vision techniques, to automatically detect and segment blue-white veil areas in dermoscopy images. The proposed method is an attempt to mimic the human perception of lesion colours, and improves and outperforms the state-of-the-art as shown in our experiments.

1 Introduction

Early detection of melanoma is paramount to patients' prognosis towards greater survival. The challenges involved in clinical diagnosis of early melanoma have provoked increased interest in computer-aided diagnosis systems through automatic analysis of skin lesion images.

Dermoscopy is a popular imaging technique, widely used by clinicians for examining and monitoring pigmented skin lesions [1]. "Inevitably, most efforts in computerizing diagnosis of melanoma lean towards automatic analysis of dermoscopy images" [2]. Dermoscopy unveils visual features in pigmented lesions that are not discernible by the naked eye.

Among dermoscopic features, colour has a substantial role. For instance, one of the most significant indicators of malignant melanoma is the blue-white veil (a.k.a. blue-grey veil) feature (with sensitivity of 51% and a specificity of 97% [3]).

Fig. 1-(a,c) show dermoscopy images of melanoma with blue-white veil feature; the blue-white veil regions are structureless areas of confluent blue pigment with a ground-glass haze (as if the image were out of focus there) [4].

The colour blue (under dermoscopy) indicates melanin localized within deeper parts of the skin [5]. Blue-white veil is found associated with hyperkeratinisation over dense amounts of melanin in the dermis [6]. While uniform blue-white structures may also be observed over blue nevus (benign) lesions, asymmetrical and irregular blue-white veil usually indicates malignant melanoma.

There exists extensive literature on skin lesion image analysis (see [8] for a survey), but here we focus on a study by Celebi et al. [7] which specifically targets automatic

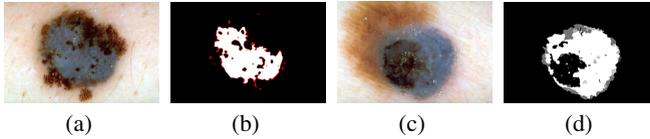


Fig. 1. (a,c): Melanoma images with Blue-white Veil; (b): veil mask by Alg. 1; The red areas are indicating disagreement between Alg. 1 and [7]; (d): veil probability map by Alg. 2

blue-white veil detection¹. Their approach involves pixel classification using explicit thresholding, where the threshold values are induced by a trained decision tree.

This paper puts forward an alternative method, by incorporating colour analysis and computer vision techniques, to automatically detect and segment blue-white veil regions in dermoscopy images. The proposed method is described next.

2 Method

We describe two approaches for automatic detection and segmentation of blue-white veil regions in dermoscopy images. The first approach is a makeover of the method [7] which performs equally well with substantively less computation. The second approach, our main contribution, is an innovative method which attempts to mimic human interpretation of lesion colours under dermoscopy. The latter incorporates colour analysis and computer vision techniques and, as will be shown, outperforms the former method in our experimental trials.

2.1 Blue-White Veil Detection by Thresholding

Celebi et al. [7] used a set of 105 images selected from [4], consisting of 43 images containing sizeable blue-white veil areas with the remaining 62 free of this feature. For each image a number of small circular regions that contain either veil or non-veil pixels were manually determined by a dermatologist and used for training.

A decision tree classifier with C4.5 [10] induction algorithm was employed to classify each pixel in the training stage into two classes: veil and non-veil. Among the 18 different colour and texture features included,² only two features appeared in the induced decision rules: The classification was conducted by thresholding on a normalized-blue channel ($B / (R + G + B)$) and relative-red feature (defined as $R - \bar{R}_s$ where \bar{R}_s is the mean of red channel values for *healthy* skin areas only).

Decision trees are simple to use and easy to understand, yet they impose disadvantages and shortcomings. For instance, their good detection rate usually arrives at the expense of high false positives. In our experiments to reproduce the method [7], using

¹ The use of blue-white veil feature has been reported in some commercially available computer-aided diagnosis (CAD) systems that do use colour information (see e.g. [9]). To our knowledge however, the only study that reports a method, experimental procedure, and results specifically pertaining to the feature being examined here is the one by Celebi et al.

² The description of features – as well as the feature extraction process – is omitted for space considerations. The interested reader is referred to [7] for details.

the training images,³ we arrived at sensitivity of 78.51% and specificity of 97.03%, which demonstrates our earlier argument about this method.

Also, caution must be taken in defining the decision boundaries (threshold values); the colour values are highly dependent on the acquisition technique and the imaging setup. For example, if illumination changes, or in cases with shadows, shading, and other colour-degrading factors, thresholding methods might fail ungracefully.

The findings of Celebi et al. indicate that blue-white veil colour data has a restricted range of values and does not fall randomly in a given colour feature space. This indicates that their method can benefit from the choice of colour representation. To investigate this, we have reproduced their training experiment where each pixel has been represented by its corresponding coordinates in RGB, nRGB, L*a*b*, YIQ, HSV, and MHG [11] colour spaces. These colour models were chosen since each represents a different family of colour spaces.

Our investigation revealed that by using normalized blue, and the L* channel (of L*a*b*) one can reproduce the results obtained by Celebi. Our finding further showed that we can replace L* with $Lum = R + G + B$ and obtain equally good results. The latter colour feature is considerably faster to compute, particularly to replace relative-red feature in the original method. Note that the computation of the relative-red feature requires a search for normal (healthy) skin colour in the background. This search also constrains the method's performance subject to the accuracy of the lesion border detection algorithm and skin colour filter used.

We further investigated the efficacy of employing alternative classification techniques. In our trials, we used Single Gaussian, Mixture of Gaussian, Bayesian, Logistic Regression, and support vector machine (SVM). Unsurprisingly, the SVM produced the best results, however, the improvements were negligible with respect to added computational cost and complexity (data not shown for space considerations).

In conclusion, we first propose a revised thresholding-based method (Algorithm 1) which is based on but improves on Celebi's approach (in term of computational cost and complexity). Testing Algorithm 1 on training images, we achieved sensitivity of 78.22% and specificity of 96.76%, which is very similar to the performance of [7]. Fig. 1-(b) illustrates the veil mask generated by this method.

Algorithm 1. – Blue-white veil detection by thresholding

```

1: Load a skin lesion image
2: for each pixel do
3:    $Lum = R + G + B$ 
4:    $nB = B/Lum$ 
5:   if  $nB \geq 0.3$  and  $0.6 \leq Lum \leq 2$  then
6:     Classify pixel as veil
7:   else
8:     Classify pixel as non-veil
9:   end if
10: end for

```

³ To conduct the experiments the entire training image-set were used, whereas training has been done on the manually selected regions as described.

An important insight to Celebi's study is that texture features were found irrelevant to detection of blue-white veil data. Celebi et al. used the classical statistical texture measures of [12] (entropy, contrast, correlation) derived from the grey level co-occurrence matrix (GLCM) of an image. We further investigated this by using two more popular descriptors in the texture classification literature: LBP [13] and MR8 [14]. Our investigation agrees with the findings in [7], attesting that texture descriptors are not discriminating features for detection of blue-white veil structure.

2.2 Blue-White Veil Detection by Discrete Colour Matching

In this section we intend to analyse the problem under study, from a further level of abstraction. That is, going beyond describing a given colour by merely indicating the location of the corresponding colours in the considered colour space.

The scheme followed in Alg. 1 and also in [7] and many others is entirely based on data analysis: it is meant for computational convenience and not to model human perception and interpretation of colour. We propose an alternative approach, as described shortly, to mimic the colour assessment performed by human observers. The alternative is an attempt to find a more perceptually and semantically meaningful way of describing colours under dermoscopy. To this aim, we ask the question of how a dermatologist identifies the presence of certain colours under dermoscopy.

Colour *naming* is to some degree subjective. However we can safely say that observers are influenced by the colours they saw previously. Indeed, a dermatologist needs training to be able to identify the blue-white veil feature. Thus, we propose to identify this feature by colour matching to a *discrete* set of colours best describing "blue-white veil". Here, we are interested in asking the question how we would name the blue-white veil colour and which colours are in this colour family.

To answer the latter questions, we analysed the veil data by mapping their colour values to the Munsell colour system [15]. In colorimetry, the Munsell colour system is one of the most fundamental colour-modellings. The Munsell system offers both perceptual and quantitative colour definitions. The perceptual definition (given in the form of a book with printed colour patches) is appropriate for the use of artists such as painters and designers, whereas the quantitative definition provides measurement standards that are appropriate for technical and scientific use.

Quantitatively, a colour in Munsell space is defined by its hue, value (lightness), and chroma (colour purity). Hue is given by a number between 0 and 10 which denotes the five principal hues: Red, Yellow, Green, Blue, and Purple, along with 5 intermediate hues (e.g., YR) halfway between adjacent principal hues. Each of these 10 hues is then broken into 10 sub-steps, so there are a total of 100 hues with integer prefixes. Value, indicating how light a colour is, is a number between 0 (signifying black) and 10 (signifying white). For any given Munsell hue and value, chroma starts at 0 (grey) and extends to a variable positive number, which represents the most saturated colour of that hue and value.

Since there is no direct conversion from standard RGB to Munsell colour quantities, we consider finding the best (closest) match between each pixel colour triple and the equivalent colour value representation of each Munsell colour patch. For colour matching, we consider working in the $L^*a^*b^*$ colour space (a.k.a. CIELAB) which is an *approximately* perceptually uniform colour model. That is, colours are organized in



Fig. 2. Colour Palette of most frequent colours identified as blue-white veil

such a way that colour differences agree more consistently with human visual perception. Thus the difference between two colours can be measured using e.g. the Euclidean distance metric. Note that the common sRGB colour space does not yield this property. In addition, CIE colour spaces are device independent, which make them suitable for colour matching and colour comparison.

Moreover, to achieve the best generalization, each colour pixel value is replaced with the colour value of its ‘superpixel’ representation. “Superpixels group pixels into perceptually meaningful atomic regions, which can be used to replace the rigid structure of the pixel grid” [16]. Superpixels also preserve information over scales and sampling resolutions, which captures image redundancy.⁴

Next, we match each colour of veil data to its closest colour patch in the Munsell system using the nearest neighbour (NN) searching technique. Interestingly, the 146,353 pixels under analysis mapped to only 116 of the totality of 2352 Munsell colour patches available in our look-up table.⁵ Among these, 98% of the veil data is described by only 80 colour patches. Fig. 2 shows these 80 colours organized on a palette according to their frequency (in a descending fashion, from left to right).

We also analysed non-veil data by the same principle. The 254,739 pixels from non-veil areas mapped to 129 Munsell colour patches, among which only 3 patches were overlapping with the 116 veil patches. These 3 contribute (all together) to less than 2% of veil data and are not considered among the 80 patches in the blue-white veil colour palette.

The colour palette (Fig. 2) can be used (much like dermatologist prior knowledge) to extract blue-white veil regions from an input image according to a nearest neighbour approach.⁶ From the training data we can also define the posterior probability of a pixel being veil given its colour in terms of the likelihood of observing its associated colour patch given the class label and the prior probability of classes.

Matching each pixel of each image in the fashion described above could be computationally expensive. To reduce this cost, we can segment colours in any given image and instead match the colour vector of the centroid of each cluster (segment) to colours in the colour palette. By colour segmentation, we decompose an image into visually homogeneous regions and effectively reduce the number of basic colour regions, while preserving salient features of the overall image. This is particularly of interest to us because in dermoscopy images there are a handful of distinct colour regions (such as

⁴ Here, we compute superpixels according to the SLIC [16] algorithm. SLIC is a simple and efficient method based on a spatially localized version of k-means clustering.

⁵ The look-up table is available at <http://www.wallkillcolor.com/mutable.csv>

⁶ The use of a colour palette is in part inspired by the work of Seidenari et al. [17].

blue-white veil) where the presence of a feature significantly affects the diagnosis while the information within such a region is often less important and can be neglected.

Colour image segmentation can be done in an automated fashion using e.g. the mean-shift method. Mean shift (MS) is a fast and robust mode-seeking algorithm; it has been used for unsupervised image segmentation where the number and the shape of the image clusters are unknown. In this study we use the EDISON [18] software.

Algorithm 2 summarizes the method described here. Testing Alg. 2 on training images, we achieved sensitivity of 83.36% and specificity of 98.04%, which outperforms Alg. 1 as well as the method [7]. Fig. 1-(d) illustrates the veil mask generated by this method (note the probabilistic [non-binary] output).

Algorithm 2. – Blue-white veil detection by discrete colour matching

```

1: Load a skin lesion image
2: Convert from sRGB to CIELAB (Assumption: light is D65)
3: Segment using EDISON (parameters are SpatialBandWidth=7, RangeBandWidth=6.5, and
   MinimumRegionArea=0.01*nRows*nColumns.)
4: for each segmented region do
5:   Find the best match from colour palette
6:   if The best match is within the tolerance (threshold) distance then
7:     Classify region as veil
8:     Assign  $P(\text{Veil}|\text{segmentcolour}) = P(\text{colourpatch}|\text{Veil}) \times P(\text{Veil})$ 
9:   else
10:    Classify region as non-veil
11:    Assign  $P(\text{Veil}|\text{segmentcolour}) = 0$ 
12:   end if
13: end for

```

3 Experiments

We tested our proposed methods on a set of 223 images selected from [4] by Celebi et al. [7]. This set consists of 173 images containing blue-white veil areas and a remaining 50 free of this feature. For this set of images, since the lesion border is provided with the ground truth, we can run the method [7] and compare the classification results, as presented in Table 1. Note that, as claimed, the proposed method Alg. 1 has produced reasonably similar results to [7] whereas it is substantively easier to compute. Moreover, the proposed method Alg. 2 performs best and in particular has higher recall (sensitivity), which is of significant importance for clinical examinations.

Accurate detection and segmentation of blue-white veil feature can be useful for computer analysis of skin lesion images. In clinical assessments however, presence or absence of this feature is associated with diagnosis. Accordingly, in a different experiment we aimed to determine only the presence (or absence) of veil feature in a set

Table 1. Classification results – Proposed methods vs. [7]

Method	n	Precision	Recall	F-Score	Specificity
Celebi et al. [7]	223	0.70	0.65	0.68	0.97
Proposed Method 1	223	0.67	0.65	0.66	0.97
Proposed Method 2	223	0.70	0.71	0.70	0.97

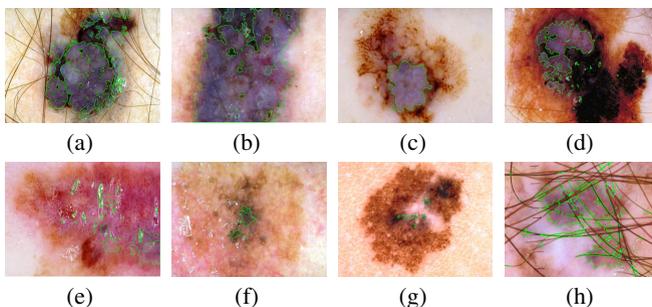


Fig. 3. (a-d): Easy images; (e-h): challenging images – The green border indicates veil areas detected by Algorithm 2. Note, the binary mask here is created by setting a threshold on the distance metric. The threshold was found empirically from training data (and set to 100 units).

of 300 images taken from various sources. The image set is divided to two subsets of 200 ‘easy’ and 100 ‘challenging’ images. An image is considered challenging if the blue-white veil area was too small, too pale, occluded, or had variegated colour.

Our experiment produced accuracy of 75% and 58% for Alg. 1 on easy and challenging sets respectively. Using Alg. 2 these results were boosted to 87% and 67% accordingly. Note that, in order to exclude very small areas which are without clinical relevance, a minimum value⁷ for areas of veil region was considered. Fig. 3 illustrates the output of Alg. 2 on some of the images in each set.

It is to be noted that the proposed method can easily be extended to account for detecting other colour features in dermoscopy images. In fact, it can be seen as a framework for colour based detection and assessment problems of similar nature. Common colours under dermoscopy are light brown, dark brown, black, blue, blue-grey, red, yellow, and white [5]. Since it is not always easy to visually distinguish these, our method (once extended to detect these *common* colours) can be effectively used as a colour clustering and colour naming engine.

The blue-white veil colour palette can be used for training purposes, as well as a reference tool (such as designers’ colour chart) for dermoscopy trainees. For that matter, a colour palette of other common colours can be generated as well. Note that the proposed method can automatically generate colour palettes given labelled training data. The advantage of this is twofold since statistical data can be extracted from training data and associated with colour patches on colour palettes. In fact, one can use this method to extract a colour map of lesions with common colours (under dermoscopy) and link it to statistical data to e.g. associate a feature with a diagnosis.

Finally, this method can aid clinical and laboratory investigations to e.g. confirm the high diagnostic relevance of presence or absence of colour and colour related features.

4 Conclusion

We have presented two schemas for automatic detection of blue-white veil feature in dermoscopy images. We first propose a revised thresholding-based method with results

⁷ Threshold is set to 0.5% of image area.

comparable to state of the art, with much reduced computation. The second approach, our main contribution, sets out an innovative method that attempts to mimic human interpretation of lesion colours. The latter outperforms prior art and moreover introduces a perceptually and semantically meaningful new approach which can serve as a scaffolding for new colour investigations in dermoscopy, for example for detection and recognition of *common* colours under dermoscopy.

Future work includes extending the proposed method to detect other *common* colours under dermoscopy, and demonstrating the usefulness of the feature under study in detection of early melanoma in a computerized automatic fashion.

References

1. Kaminska-Winciorek, G., Spiewak, R.: Tips and tricks in the dermoscopy of pigmented lesions. *BMC Dermatology* 12(1) (August 2012)
2. Madooei, A., Drew, M.S., Sadeghi, M., Atkins, M.S.: Automated pre-processing method for dermoscopic images and its application to pigmented skin lesion segmentation. In: 20th Color and Imaging Conference, pp. 158–163 (2012)
3. Menzies, S.: *Atlas of Surface Microscopy of Pigmented Skin Lesions*. McGraw Hill (2003)
4. Argenziano, G., et al.: *Interactive atlas of dermoscopy*. Edra Medical Publishing & New Media, Milan (2000)
5. Soyer, H.P., et al.: Dermoscopy of pigmented skin lesions. *EJD* 11(3), 270–276 (2001)
6. Adams, J.: Dermoscopic blue hue and blue-white veil in skin lesions indicate different diagnoses. *Am. J. Dermatopathol.* (23), 463–469 (2001)
7. Celebi, M.E., et al.: Automatic detection of blue-white veil and related structures in dermoscopy images. *Computerized Medical Imaging and Graphics* 32(8), 670–677 (2008)
8. Korotkov, K., Garcia, R.: Computerized analysis of pigmented skin lesions: a review. *Artificial Intelligence in Medicine* 56(2), 69–90 (2012); PMID: 23063256
9. Menzies, S., et al.: The performance of SolarScan. *Archives of Dermatology* (11), 1388–1396 (2005); PMID: 16301386
10. Quinlan, J.R.: *C4. 5: programs for machine learning*, vol. 1. Morgan Kaufmann (1993)
11. Madooei, A., Drew, M.S., Sadeghi, M., Atkins, M.S.: Intrinsic melanin and hemoglobin colour components for skin lesion malignancy detection. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) *MICCAI 2012, Part I. LNCS*, vol. 7510, pp. 315–322. Springer, Heidelberg (2012)
12. Haralick, R., Shapiro, L.: *Computer and Robot Vision*, p. 459. Addison-Wesley (1992)
13. Ojala, T., et al.: Multiresolution Gray-Scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal.* 24(7), 971–987 (2002)
14. Varma, M., Zisserman, A.: A statistical approach to texture classification from single images. *Int. J. of Computer Vision* 62(1), 61–81 (2005)
15. Landa, E.R., Fairchild, M.D.: Charting color from the eye of the beholder. *American Scientist* 93(5), 436–443 (2005)
16. Achanta, R., et al.: SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* 34(11), 2274–2282 (2012)
17. Seidenari, S., Pellacani, G., Grana, C.: Computer description of colours in dermoscopic melanocytic lesion images reproducing clinical assessment. *British J. of Dermatology* 149(3), 523–529 (2003)
18. Georgescu, B., Christoudias, C.M.: The edge detection and image segmentation (EDISON) system (2003), Code available at <http://coewww.rutgers.edu/riul/research/code/EDISON/>