

# Highlight Detection by 2D-Histogram Analysis

Allan Hanbury  
*Pattern Recognition and Image Processing Group*  
*Vienna University of Technology, Favoritenstraße 9/1832*  
*A-1040 Vienna, Austria*

## Abstract

We present image segmentation and highlight detection algorithms based on the dichromatic reflection model. For image segmentation, we use the model prediction that objects of a certain colour produce lines (the matte lines) radiating away from the origin of the RGB colour space. These lines therefore show up as peaks in a 2-dimensional histogram of the angular coordinates of a spherical polar coordinate representation of the RGB space. An algorithm for automatically locating these peaks is suggested. When the matte line locations are known, one can define cylindrical polar coordinate systems having their  $z$ -axes centred on the matte lines. We suggest a Hough-based algorithm for the detection of highlight lines in 2D-histograms of the  $\rho$  and  $z$  coordinates of the cylindrical polar coordinate system. Examples of the results of applying these algorithms are given.

## Introduction

Physics-based methods, such as those based on the *dichromatic reflection model* for dielectric materials [1], have been applied to the segmentation and analysis of colour images based on their RGB histograms, as well as to the estimation of scene properties such as surface roughness and illuminant colour [2]. In general, the use of physics-based techniques has been limited to images captured under carefully controlled conditions which contain only objects which have physical characteristics described by the model. It is thus known in advance that the colour histogram will have a form which is convenient to process. Klinker et al. [1], for example, describe a rather complex pixel clustering algorithm based on the dichromatic reflection model, which is demonstrated to work well on an image of a group of plastic objects. We investigate in this paper the applicability of the dichromatic reflection model to the analysis of arbitrary images about which no prior information is known. Such images are found in abundance in digital photograph collections and on the Internet, so it is important to determine how much can be done by analysing them based on the assumptions of a relatively simple physical model.

The dichromatic reflection model predicts that when

a ray of light interacts with a material, it splits into two. One part is directly reflected by the surface and the other part enters the material. Some of the light that enters the material will eventually, through scattering and refraction by the material, exit through the surface by which it entered. These two processes are referred to as *surface reflection* and *body reflection*. The surface-reflected light causes highlights (specular reflection) on the material. In the RGB space, the model predicts that the colour of an object satisfying the model will make up two linear clusters. The *matte line*, corresponding to body reflection, passes through the origin (as long as there is no ambient illumination in the scene [2]). The *highlight line* corresponds to surface reflection, and extends away from the matte line somewhere along its length. The two clusters together are therefore in the shape of a skewed-T or skewed-L. In the more general case of diffuse surface reflection, it could also be in the shape of a skewed-P. Dielectric materials include porcelain (ceramic), glass, plastics, and the oxides of various metals. More specifically, Tominaga [3] has shown experimentally that plastics, paints, ceramics, vinyls, tiles, fruits, leaves and woods are well described by the dichromatic reflection model; whereas metal, cloth and paper are not. This list covers a large proportion of materials which are often present in images, and so supports our application of this model to such general images. However, the developed algorithm is not guaranteed to work in all cases — sometimes the assumptions are violated to too great an extent. It will have to be combined with other techniques and possibly other physical models to produce a completely general algorithm.

Recently there has been some work done on the automated detection of highlights in images through the use of 2D histograms of the brightness and saturation coordinates [4, 5] in a 3D-polar coordinate colour representation. In these histograms, the colours of pixels in highlight regions tend to form straight lines in the high brightness part of the histogram. The highlight lines in the RGB space are effectively all projected onto a plane in the histogram. What is not mentioned is that this method is particularly good for detecting *colour clipped* pixels. These are the regions of a scene whose brightness exceeds the dynamic range of the camera. The pixels corresponding to these regions tend to lie along the faces of the RGB cube [1].

This implies that in the brightness-saturation histogram, they will lie along the line corresponding to the maximum saturation for a given brightness, the position of which can easily be determined. Torres et al. [5] take the rather brutal approach of performing a histogram equalisation of the brightness channel before calculating the 2D histogram. While the results demonstrated look good, this step should guarantee the detection of some highlights in every image, even if none are present.

Our aim is to develop a segmentation and highlight detection algorithm based on characteristics of the dichromatic reflection model and to demonstrate that it works well on a number of real images. Ong and Matsuyama [6] have suggested a clustering algorithm in the RGB space, but their experimental results are limited to images containing only one or two materials. The Markov Random Field approach based on the dichromatic reflection model proposed in [7] is only tested on a single artificial image. In our approach, we first make use of the property that all the matte lines intersect the origin of the RGB space. A representation of RGB coordinates in terms of spherical polar coordinates therefore allows these lines to be straightforwardly detected in a 2D-histogram. Once we have the position of the matte lines, colour and highlight analysis for each matte line are done by representing the RGB coordinates in terms of cylindrical polar coordinates centred on each matte line. We examine the use of 2D histogram analysis in this coordinate system for the detection of highlights in images.

## Segmentation in Spherical Polar Coordinates

As our goal is the segmentation of arbitrary images, we attempt to use the dichromatic reflection model to as great an extent as possible, while keeping the segmentation algorithms simple. We begin by demonstrating the use of a spherical polar coordinate system for highlighting the linear matte clusters, and then describe an algorithm for automatically segmenting the images based on their 2-dimensional histograms in spherical coordinates.

### Spherical polar coordinates

As a demonstration of the clustering of colours into linear structures, we plot, in Figure 1b, the RGB histogram of Figure 1a, for which no special effort has been made to include only objects satisfying the dichromatic reflection model. The points in this histogram have colours given by their corresponding RGB coordinates. One can see in the histogram that there are linear clusters corresponding to the blue slide and to the skin and a (less-linear) cluster corresponding to the grass. Two artefacts due to the acquisition procedure are also visible:

- The blue cluster changes direction and lies along a face of the RGB cube in the high brightness region

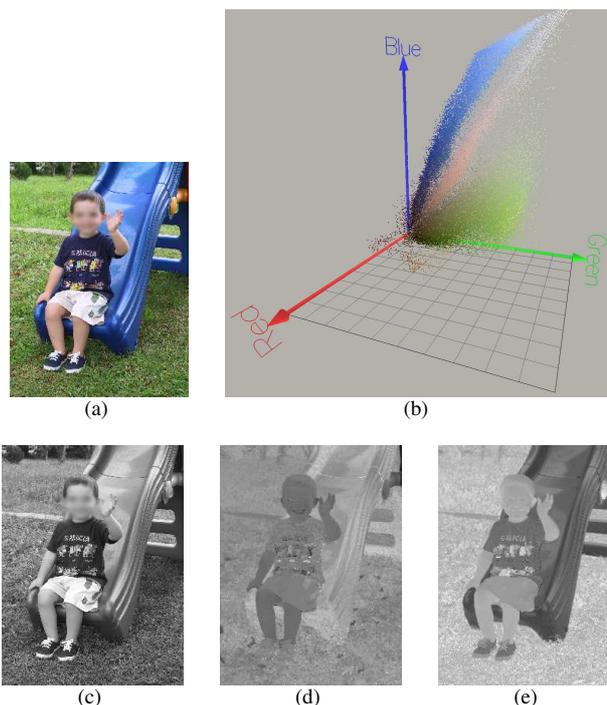


Figure 1: (a) Initial image. (b) RGB histogram of (a) (created using the Colorspace software [10]). The (c)  $M$ , (d)  $\theta$  and (e)  $\phi$  spherical coordinates of image (a)

of the RGB space. This *colour clipping* [8] is due to the limited dynamic range of the camera.

- The clusters curve slightly. This is due to the gamma-correction of the camera [8]. It could potentially be corrected by one of the gamma-correction compensation algorithms available [9].

As we assume that the linear clusters radiate out from the origin, a representation of the RGB values in spherical coordinates should result in pixels belonging to the same linear cluster having similar angular coordinates. Bajcsy et al. [11] have used a similar approach in their  $S$  space. The disadvantage of the  $S$  space in the present context is that it requires the scene illumination colour to be known. While algorithms to estimate this exist [12], they introduce extra uncertainty into the process. The conversion from RGB to spherical coordinates is done as follows:

$$M = \sqrt{R^2 + G^2 + B^2} \quad (1)$$

$$\theta = \tan^{-1} \left( \frac{G}{R} \right), \quad \phi = \cos^{-1} \left( \frac{B}{M} \right) \quad (2)$$

The images corresponding to the spherical coordinate representation of Figure 1a are shown in Figure 1c–e. Notice that for the  $\theta$  and  $\phi$  images, the coordinates corresponding to a specific object are rather uniform. This can be better seen by looking at a two-dimensional histogram of the  $\theta$  and  $\phi$  coordinates, shown in Figure 2. On this histogram, the clusters corresponding to various objects in the initial image have been manually annotated.

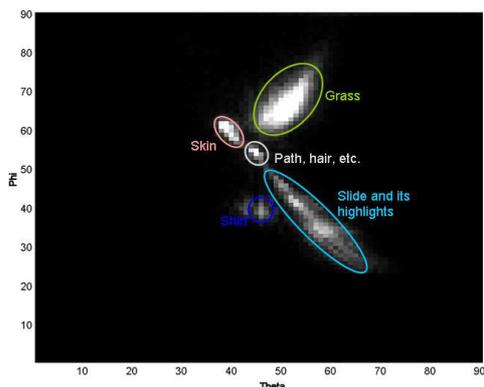


Figure 2: The 2-dimensional  $\theta$ - $\phi$  histogram. The regions to which the clusters correspond have been entered by hand.

### Segmentation based on the 2D histogram

Based on the information visible in the 2D  $\theta$ - $\phi$  histogram, it is simple to produce a segmentation of the initial image. One places straight lines in the RGB space passing through the origin in the direction of each detected cluster in the histogram (an algorithm for automatic cluster detection is presented later). For each pixel in the initial image, the line closest to it in the RGB space is found, and it is assigned to the cluster described by that line. In 3D space, the shortest distance  $d$  between a line passing through points  $\mathbf{x}_1 = (x_1, y_1, z_1)$  and  $\mathbf{x}_2 = (x_2, y_2, z_2)$ , and a point  $\mathbf{x}_0 = (x_0, y_0, z_0)$  is given by

$$d = \frac{|(\mathbf{x}_2 - \mathbf{x}_1) \times (\mathbf{x}_1 - \mathbf{x}_0)|}{|\mathbf{x}_2 - \mathbf{x}_1|} \quad (3)$$

where  $\times$  indicates the cross product between two vectors.

To segment a colour image described by the function  $f(\mathbf{x})$  giving the vector  $(R, G, B)$  at position  $\mathbf{x}$  in the image, the following algorithm is used:

1. The cluster centres in the 2D  $\theta$ - $\phi$  histogram are found, either manually or automatically, giving  $N$  pairs of coordinates  $\mathbf{L}_i = (\theta_i, \phi_i)$ ,  $i = 1, 2, \dots, N$ .
2. The coordinates  $\mathbf{L}_i$  are converted to RGB coordinates  $\mathbf{L}_i^C = (R_i, G_i, B_i)$ . The value of  $M$  in this conversion is arbitrary, and is taken to be 1.
3. For every pixel  $\mathbf{x}$  in the input colour image:
  - (a) Equation 3 is evaluated for each  $\mathbf{L}_i^C$ , taking  $\mathbf{x}_1 = (0, 0, 0)$ ,  $\mathbf{x}_2 = \mathbf{L}_i^C$  and  $\mathbf{x}_0 = f(\mathbf{x})$ , producing  $N$  distances  $d_i$ .
  - (b) The pixel is assigned to the cluster  $j$  corresponding to the smallest distance, i.e.  $d_j = \min_i (d_i)$ .

### Automatic segmentation of the 2D histogram

We first consider histogram creation. One of the histogram parameters which can have an effect on the automatic extraction of maxima is the quantisation. It was experimentally determined that a histogram quantisation step of  $2^\circ$

worked well with the proposed algorithm. This larger quantisation step has the effect of reducing the number of local maxima in the histogram.

Colour clipping in the histogram can cause some of the blobs in the histogram to spread out, and hence can perturb the finding of blob centres. We therefore implemented a heuristic solution to this problem proposed by Klinker et al. [8]. Potentially colour clipped pixels can be identified by placing a high threshold on the values of the three colour channels. Pixels which have a value in one of the  $R$ ,  $G$  or  $B$  channels greater than a threshold (here 240) are excluded from the histogram. The histogram of Figure 1a calculated using these two modifications is shown in Figure 3a. It can be processed as a greylevel image.

An important step is automatically finding the blob centres. They are initially found by using the  $h$ -maxima operator [13]. The parameter  $h$  indicates how high a peak should be with respect to the surrounding region to qualify as a valid maximum. We found experimentally that a value of  $5 \times 10^{-4}$  multiplied by the number of pixels in the image under consideration works well. The results are not sensitive to small changes in the value of this parameter as the peaks are generally well-defined. The positions of the five maxima found by applying the  $h$ -maxima operator to the histogram of Figure 3a are shown in Figure 3b. Note that the two points at the lower right are a single maximum, as 8-connectivity was used.

The next step is to refine the positions of the maxima by calculating a greylevel-weighted centre of mass of each blob. Before this can be done, we need to find the “zones of influence” of each maximum, which we do by using the watershed operator [13]. This is applied to the inverted histogram (the flooding starts from the minima) on which the maxima found by the  $h$ -maxima operator have been imposed as minima. Figure 3c shows the watershed lines and regions found by the watershed operator. The greylevel-weighted centre of mass is now calculated within each region on the histogram in the standard way.

The centres of mass of the five regions located by the algorithm are given by the following  $(\theta, \phi)$  coordinates: (44.9, 40.3), (44.9, 53.3), (39.1, 61.1), (51.7, 68.4) and (59.7, 34.4). The segmentation obtained when using these cluster centres in the process described previously is shown in Figure 5b, in which each cluster is indicated in a different randomly-chosen colour. In general, the segmentation corresponds well to the objects in the image. Almost all regions which are misclassified correspond to highlights, such as those on the slide and on the boy’s forehead. This is understandable, as the highlights give rise to side branches of the matte clusters, which have not yet been taken into account. Other algorithms, such as graph-theoretical clustering [14], could also be applied to this cluster-finding task. A further example of such a segmentation applied to Figure 7a is shown in Figure 7b. Here again, the five regions detected correspond well to actual objects in the scene. More examples are given in [15].

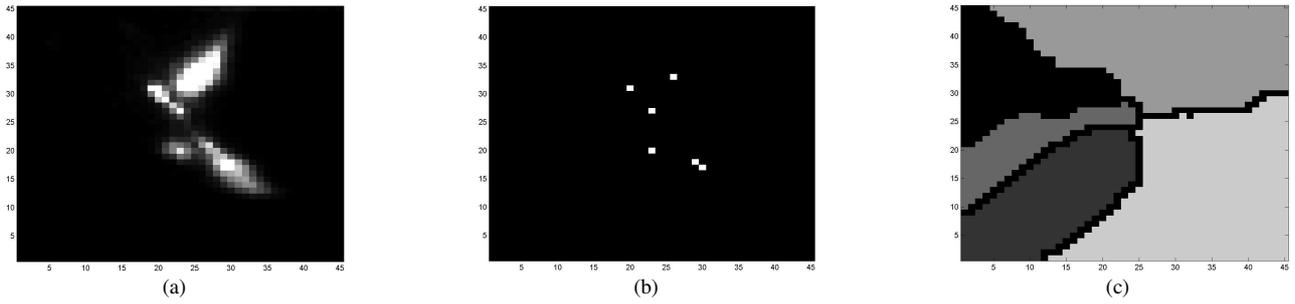


Figure 3: (a) The  $\theta$ - $\phi$  histogram of Figure 1a with bins of size  $2^\circ$ . The values on the axes should be multiplied by 2 to get values in degrees. (b) The  $h$ -maxima of histogram (a). (c) The watershed of the inversion of (a) on which (b) have been imposed as minima.

### Detecting highlights

We now use the knowledge of the position of a matte line extracted during the segmentation step to attempt to locate highlight lines radiating out from it. The crucial step is the conversion to a cylindrical polar coordinate system with its  $z$ -axis centred on the matte line to be analysed. Image analysis methods developed for use in cylindrical polar coordinate colour spaces (HSV, HLS, etc.) are also applicable in this representation.

#### Cylindrical polar coordinates

To align the  $z$ -axis of the cylindrical polar coordinate system with the matte line having a direction given by the spherical coordinates  $\theta$  and  $\phi$ , we rotate the  $RGB$  axes so that the  $B$  axis is aligned with the matte line before converting to cylindrical coordinates. This is done by multiplying each  $RGB$  vector by the following rotation matrix:

$$\begin{bmatrix} \cos(\phi) \cos(-\theta) & -\cos(\phi) \sin(-\theta) & \sin(\phi) \\ \sin(-\theta) & \cos(-\theta) & 0 \\ -\sin(\phi) \cos(-\theta) & \sin(\phi) \sin(-\theta) & \cos(\phi) \end{bmatrix} \quad (4)$$

The conversion to  $\rho$ ,  $\theta_c$  and  $z$  cylindrical coordinates is then done as follows:

$$\rho = \sqrt{R_r^2 + G_r^2}, \quad \theta_c = \tan^{-1} \left( \frac{G_r}{R_r} \right), \quad z = B_r \quad (5)$$

where  $R_r$ ,  $G_r$  and  $B_r$  are the rotated  $R$ ,  $G$  and  $B$  coordinates. Note that  $\theta$  is an angular coordinate in the spherical coordinate system, and  $\theta_c$  in the cylindrical coordinate system. In the example image, the spherical coordinates of the blob corresponding to the slide are  $\theta = 59.7$  and  $\phi = 34.4$ . The images showing the  $\rho$ ,  $\theta_c$  and  $z$  coordinates when the  $z$ -axis is aligned with the matte line having these coordinates are shown in Figure 4<sup>1</sup>. In this representation, the pixels having a colour closest to the selected matte line have the smallest  $\rho$  values, as can be seen

<sup>1</sup>If one substitutes spherical coordinates  $\theta = 45^\circ$  and  $\phi = 54.7^\circ$  into the rotation matrix, the  $B$ -axis lines up with the achromatic axis (i.e. the axis through all the grey colours), giving the standard cylindrical polar coordinate colour representation.

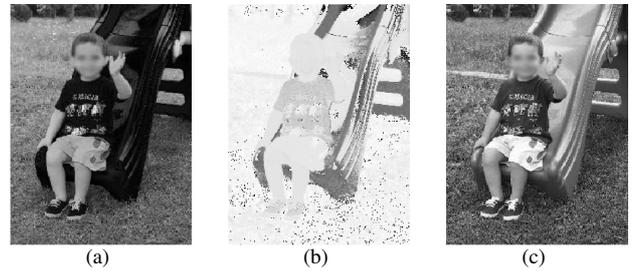


Figure 4: The (a)  $\rho$ , (b)  $\theta_c$  and (c)  $z$  coordinates of the cylindrical polar coordinate representation with the  $z$ -axis aligned with the matte line corresponding to the slide.

for the slide region. Highlight lines extending away from the matte line will be characterised by increasing  $\rho$  values and relatively constant  $\theta_c$  values. As the highlights are bright, they should also have high  $z$  values. It can be seen that the reflections on the slide have these characteristics. Two possible approaches are conceivable: an analysis of the images to find, for example, areas adjacent to the slide which have the required characteristics to be highlights, as done in [11], or the analysis of histograms to find potential highlight pixels. We pursue the latter approach.

#### Highlight line detection

To find the highlight line, we begin by calculating a one-dimensional  $\theta_c$  histogram weighted by the  $z$ -values (the bins are incremented by the  $z$ -value corresponding to each  $\theta_c$ -value instead of by 1). The  $z$ -weighting is used as it is expected that the highlights are bright and therefore have high  $z$  values. The highest peak in the histogram, at position  $\theta_{\max}$ , is the expected position of the highlight line. The  $z$ -weighted  $\theta_c$  histogram calculated using the cylindrical coordinate system centred on the slide is shown in figure 5a. The highest peak in this histogram is at  $327^\circ$ . At present, we only take the position of the highest peak into account in further steps.

Figure 6a shows the 2D histogram having  $z$ -values on the horizontal axis and  $\rho$ -values on the vertical axis for the cylindrical coordinate system centred on the slide.

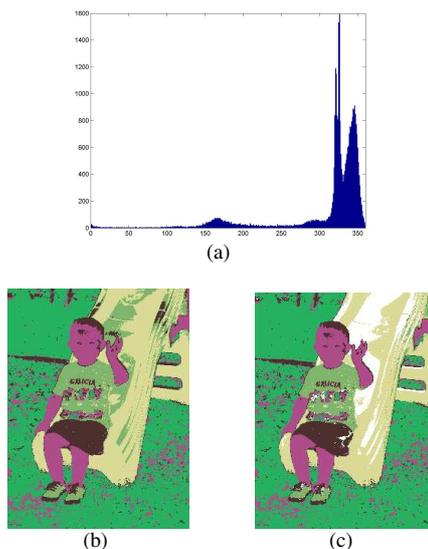


Figure 5: (a) The  $z$ -weighted  $\theta_c$  histogram. (b) Figure 1a segmented by the suggested algorithm. (c) The detected highlights marked on (b).

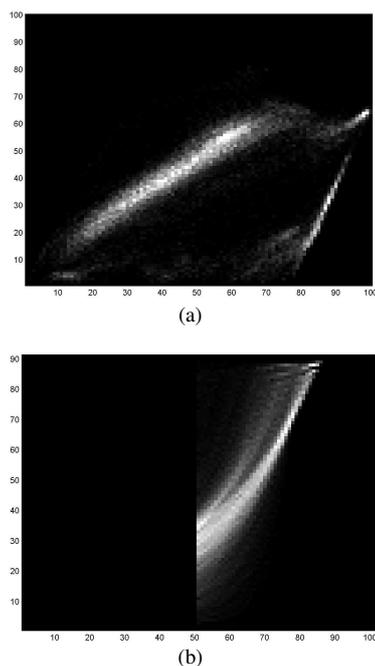


Figure 6: (a) The  $z$ - $\rho$  2D histogram for pixels in the angular range  $\theta_c = 327 \pm 20^\circ$  for the images in Figure 4. (b) The Hough-type transform of histogram (a).

The highlight line is clearly visible on the right, starting from the horizontal axis. The cluster further up is another matte line. This histogram includes only those pixels which have  $\theta_c$  values in the range of  $20^\circ$  on either side of  $\theta_{\max}$ , i.e.  $327^\circ \pm 20^\circ$ . This limit prevents the highlight line from being masked by extraneous information due to other colours in the image.

Based on the underlying reflection model, we expect the highlight lines to: lie in the high brightness area of the histogram, start close to the matte line (which in this histogram is the horizontal axis), and slope toward the right in the histogram. These lines are therefore most usefully characterised by the following two parameters:  $z_l$ , the starting position of the line on the horizontal axis and  $\theta_l$ , the angle of the line with respect to the horizontal axis (measured between the positive direction of the horizontal axis and the line). We use a Hough transform-based algorithm to find the coordinates of the highlight line accurately. We incorporate the first two assumptions about the position of the highlight line by limiting the search to the region of the histogram satisfying  $\rho \leq 0.25$  and  $z \geq 0.5$ . For all pixels lying in this region of the histogram, lines having all possible values of  $z_l$  and  $\theta_l$  are fitted, and their coordinates in a Hough accumulator are incremented. This accumulator calculated for the histogram in Figure 6a is shown in Figure 6b, where the  $z_l$  coordinates are on the horizontal axis, and the  $\theta_l$  coordinates on the vertical axis. The maximum value in this histogram is at  $z_l = 77$  and  $\theta_l = 69^\circ$ , which corresponds well to the highlight line visible on the corresponding 2D-histogram.

For visualisation, we mark all pixels having colour coordinates lying within the cylinder with radius 0.15 centred on the detected highlight line and having  $\rho$ -coordinates between 0.05 and 0.5 as white on the segmented image, shown in Figure 5c. It is clear that the highlights on the slide are extremely well detected, with those on the steps behind the slide detected too.

A further example of highlight detection for the skin region of Figure 7a is shown. The 2D  $\rho$ - $z$  histogram centered on the matte line corresponding to the skin is shown in Figure 7c, in which it is visible that as the highlight region is so small, the highlight branch is extremely faint. The Hough-based algorithm fails for this particular case (it has a maximum at  $\theta_l = 86^\circ$  and  $z_l = 61$ ) as it is influenced by the background pixels. A possible solution to this problem, which remains to be tested, is to eliminate all pixels having colour coordinates lying within a specified distance from any of the matte lines (except for the one forming the centre of the coordinate system) from consideration. The highlights found using manually determined coordinates for the highlight line ( $\theta_l = 60^\circ$  and  $z_l = 75$ ) are shown in Figure 7d. The labelling of the highlight regions has not been limited to the skin areas in this image. Looking at these regions, one sees that the highlights on the faces and hand have been well detected. However many other white regions, such as the clothes, have also been marked as highlights. Further examples are presented in [15].

## Conclusion

The use of the dichromatic reflection model in segmenting and detecting highlights in arbitrary images has been

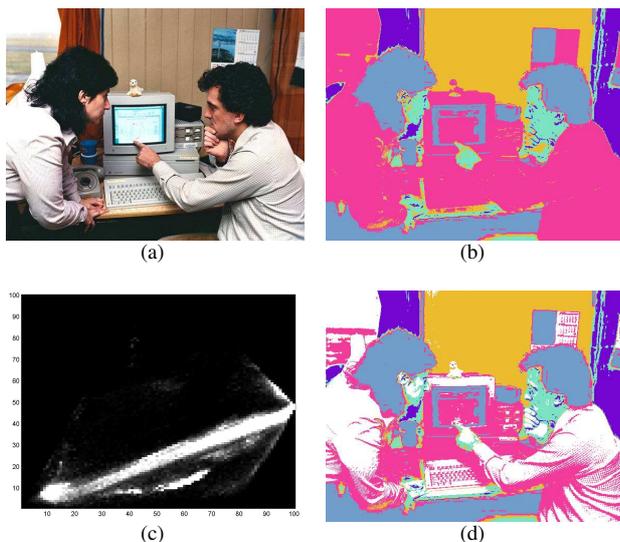


Figure 7: (a) Initial image (NASA). (b) Segmented image. (c) 2D  $\rho - z$  histogram for the direction  $\theta_c = 146^\circ$ . (d) Highlight regions for the manually determined coordinates.

discussed. For segmentation, we make use of the model prediction that all matte lines radiate out from the origin in the RGB space. This suggests that a representation of RGB coordinates in terms of spherical polar coordinates is useful for segmentation. At present, we optimise the coordinates of the matte lines found by finding the centres of the clusters in the 2D  $\theta - \phi$  histogram as accurately as possible. An improvement in accuracy is likely if the matte lines are fitted to the clusters in the RGB space once the approximate centres have been found on the histogram. A point to be kept in mind is that because all the matte lines intersect at the origin, classification of dark-coloured pixels into regions is most likely arbitrary. The  $M$  coordinate, however, can be used as an indicator of the certainty of the classification. Furthermore, if a detected cluster does not contain any pixels near the origin and does not lie along a line radiating away from the origin, this could be evidence of a non-dielectric object in the image.

Once the matte lines have been found, we suggest a further analysis in a cylindrical polar coordinate system centred on the matte lines, which is demonstrated for highlight detection. The method for the detection of highlight lines is still in need of improvement and further automation. For example, the background clutter in the  $\rho - z$  histograms could be reduced by excluding all pixels which are close to matte lines not under consideration. Further information that could be used is that all the highlight lines are parallel and in the direction of the illumination colour [8]. Methods for estimating the illumination colour [12] could potentially be used to limit the search directions for the highlight line, thereby eliminating the need to use the  $z$ -weighted  $\theta_c$  histogram. A comparison to clustering methods, such as the one described in [6], as well as the

potential complementarity of these approaches remains to be investigated.

## References

1. Gudrun J. Klinker, Steven A. Shafer and Takeo Kanade, Image segmentation and reflection analysis through color, in Applications of Artificial Intelligence VI, SPIE, vol. 937, pp. 229–244 (1988).
2. Carol L. Novak and Steven A. Shafer, Estimating scene properties from color histograms, Tech. Rep. CMU-CS-92-212, School of Computer Science, Carnegie Mellon University (1992).
3. Shoji Tominaga, Surface identification using the dichromatic reflection model, IEEE Transactions on Pattern Analysis and Machine Intelligence, 13(7), 658 (1991).
4. Jesús Angulo and Jean Serra, Colour feature extraction from luminance/saturation histogram in  $L_1$  representation, Tech. Rep. N-01/03/MM, Centre de Morphologie Mathématique, Ecole des Mines de Paris (2003).
5. Fernando Torres, Jesús Angulo and Francisco Ortiz, Automatic detection of specular reflectance in colour images using the MS diagram, in Proceedings of the CAIP, pp. 132–139 (2003).
6. Chun-Kiat Ong and Takashi Matsuyama, Robust color segmentation using the dichromatic reflection model, in Proceedings of the 14th International Conference on Pattern Recognition, vol. 1, pp. 780–784 (1998).
7. Paul Fieguth and Slawo Wesolkowski, Highlight and shading invariant color image segmentation using simulated annealing, in Proceedings of the 3rd International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition, Springer-Verlag, pp. 314–327 (2001).
8. Gudrun J. Klinker, Steven A. Shafer and Takeo Kanade, The measurement of highlights in color images, International Journal of Computer Vision, 2, 7 (1988).
9. Graham Finlayson and Ruixia Xu, Illuminant and gamma comprehensive normalisation in  $\log rgb$  space, Pattern Recognition Letters, 24, 1679 (2003).
10. Philippe Colantoni, Colorspace software, <http://www.couleur.org> (2003).
11. Ruzena Bajcsy, Sang Wook Lee and Aleš Leonardis, Detection of diffuse and specular reflections and inter-reflections by color image segmentation, International Journal of Computer Vision, 17(3), 241 (1996).
12. Graham D. Finlayson and Gerald Schaefer, Solving for colour constancy using a constrained dichromatic reflection model, International Journal of Computer Vision, 42(3), 127 (2001).
13. Pierre Soille, Morphological Image Analysis: Principles and Applications, Springer-Verlag (1999).
14. K. Fukunaga, Introduction to Statistical Pattern Recognition, Academic Press (1990).
15. Allan Hanbury, Physics-based segmentation of colour images in spherical coordinates, Tech. Rep. PRIP-TR-84, TU Wien (2003).

## Biography

Allan Hanbury received the B.Sc. (Hons.) and M.Sc. degrees in Physics from the University of Cape Town, South Africa, in 1995 and 1999, respectively, and the Ph.D. degree in Mathematical Morphology from the Paris School of Mines, France, in 2002. He is presently an assistant professor in the Pattern Recognition and Image Processing Group of the Vienna University of Technology, Austria. His research interests include mathematical morphology, colour and texture analysis and industrial applications.