 Cytometry 39:275–284 (2000)

Hue-Saturation-Density (HSD) Model for Stain Recognition in Digital Images From Transmitted Light Microscopy

Jeroen A.W.M. van der Laak,* Martin M.M. Pahlplatz, Antonius G.J.M. Hanselaar, and Peter C.M. de Wilde

Department of Pathology, University Hospital Nijmegen, Nijmegen, The Netherlands

Received 6 July 1999; Revision Received 3 December 1999; Accepted 9 December 1999

Background: Transmitted light microscopy is used in pathology to examine stained tissues. Digital image analysis is gaining importance as a means to quantify alterations in tissues. A prerequisite for accurate and reproducible quantification is the possibility to recognise stains in a standardised manner, independently of variations in the staining density.

Methods: The usefulness of three colour models was studied using data from computer simulations and experimental data from an immuno-doublestained tissue section. Direct use of the three intensities obtained by a colour camera results in the red-green-blue (RGB) model. By decoupling the intensity from the RGB data, the hue-saturation-intensity (HSI) model is obtained. However, the major part of the variation in perceived intensities in transmitted light microscopy is caused by variations in staining density. Therefore, the hue-saturation-density (HSD) transform was defined as the RGB to HSI transform, applied to optical density values rather than intensities for the individual RGB channels.

Results: In the RGB model, the mixture of chromatic and intensity information hampers standardisation of stain recognition. In the HSI model, mixtures of stains that could be distinguished from other stains in the RGB model could not be separated. The HSD model enabled all possible distinctions in a two-dimensional, standardised data space.

Conclusions: In the RGB model, standardised recognition is only possible by using complex and time-consuming algorithms. The HSI model is not suitable for stain recognition in transmitted light microscopy. The newly derived HSD model was found superior to the existing models for this purpose. Cytometry 39:275–284, 2000.

Key terms: true colour image analysis; stain recognition; image segmentation; colour models

Classically, light-absorbing stains are used in diagnostic pathology to visualise cells and cell constituents in cytological and histological specimens. Moreover, use of multiple contrasting stains in a single specimen enables the simultaneous visualisation of different components under a microscope (1). Quantitative analysis of images acquired by a CCD camera mounted on a microscope may be used to obtain reproducible information from such specimens (2–4). A CCD colour camera measures the intensity of transmitted light in three wavelength ranges in the visible spectrum for each point (*pixel*) in the microscopic image (5). The resulting red, green, and blue (RGB) intensities together may be regarded as a point in a three-dimensional (3D) Euclidean coordinate space (6,7). These three intensities reflect both the absorption characteristics of the used stains and the amount in which those stains are present [Lambert-Beers' law (8)]. Variations in the amount of stain bound to different locations of the specimen will result in variations in the intensities acquired for different pixels, even if only a single stain is present. Within the RGB space, these intensity variations will comprise a complexly shaped 3D domain for each stain (9).

To enable analysis of stained objects in digital images from microscopic specimens, a procedure for the classification of pixels as belonging to either background or to one of the used stains is required. Pixels originating from mixtures of stains should be classified as such. Classification should be based on the absorption characteristics of the used stains and should be independent of the amount of stain. Thus, a pixel originating from a small amount of a certain stain should be recognised in the same way as a pixel originating from a large amount of the same stain. Classification can be performed by determining in which of the 3D RGB subspaces, representing the used stains,

Grant sponsor: Dutch Cancer Society; Grant number: NUKC-NKB-94-726.

*Correspondence to: Jeroen A.W.M. van der Laak, M.Sc., Department of Pathology, University Hospital Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands.

E-mail: j.vanderlaak@pathol.azn.nl

each pixel in the image is located. For this, exact knowledge of the 3D subspaces is necessary. Pixels originating from a mixture of stains will be located between the domains of the separate stains in the RGB space. The disadvantage of this method of classification is that calculations involving complexly shaped 3D domains require complicated and time-consuming algorithms (9). To avoid the use of such complex algorithms, often a projection of the 3D data onto a line or a plane is used (7,9). After the projection, different stains present in the image should be well enough separated from each other to be recognised independently. For each application, the optimal projection can be determined by inspection of data resulting from a number of test specimens. Using this projection for stain recognition in a specimen with staining characteristics that deviate from those in the test set may appear to be inadequate (7). There is no single "optimal" projection, which reflects the absorption characteristics of the used stains.

Use of the RGB model is problematic because the information of interest, i.e., the colour of the stain (determined by the absorption characteristics), is mixed with variations in the amount of stain. A widely used procedure to extract the chromatic (colour) information from the RGB data is the hue-saturation-intensity (HSI) model (6). The RGB to HSI transform decouples the intensity information from the colour information (6). In the HSI colour space, classification of pixels is based on two dimensions, hue (dominant wavelength) and saturation (purity of the colour). These are independent of the signal intensity. We found that the HSI model is inadequate for the recognition of absorbing stains, because the variations between different pixels originating from a single stain are caused by variations in the amount of stain and thus show a logarithmic relationship with the camera channel intensities instead of a linear one. Therefore, the hue and saturation of the HSI model are not a correct representation of the absorption characteristics of the used stains.

In this study, we propose and test an adaptation of the HSI model in which the RGB to HSI transform is applied to optical densities (OD) for the individual RGB channels instead of intensities. The model will be called the hue-saturation-density (HSD) model. As the conversion from channel intensity to OD is a non-linear one, the resulting chromatic component of the HSD model differs from that of the HSI model. The chromatic component of the HSD model is independent of the amount of stain, and therefore offers a good representation of the absorption characteristics of the used stains. The model was evaluated both in a computer-simulated model and on microscopic images from immuno-doublestained tissue sections.

MATERIALS AND METHODS

Physical Aspects of Image Formation in Transmitted Light Microscopy

Monochromatic light travelling through an absorbing medium satisfies Lambert-Beers' law of absorption (8). This law describes an exponential relationship between the intensity of monochromatic light transmitted through a specimen and the amount of stain present in the specimen:

$$I(\lambda) = I_0(\lambda) \cdot e^{-A \cdot c(\lambda)} \tag{1}$$

where $I(\lambda)$ is the intensity of light of wavelength λ transmitted through the specimen, $I_0(\lambda)$ is the intensity of light of wavelength λ entering the specimen, A is the amount of stain per unit area of the specimen, and $c(\lambda)$ is a wavelength-dependent factor reflecting the absorption characteristics of the particular stain *(absorption curve).* The absorption curve determines the fraction of the incident light of wavelength λ transmitted through the stain.

Based on the human eye, three-chip CCD RGB cameras use three broad-band filters for colour image acquisition. These are called the red (R), green (G), and blue (B) channels, based on the perceived colours they represent. The camera sensitivity curve defines the relationship between the incident light and the electric output signal of the camera for each channel. The output of the camera for channel *cb* is given by

$$I_{cb} = \int_0^\infty S_{cb}(\lambda) \cdot I_0(\lambda) \cdot e^{-A \cdot c(\lambda)} d\lambda \tag{2}$$

where *cb* is R, G, or B and $S_{cb}(\lambda)$ is the sensitivity of channel *cb* at wavelength λ.

In this article, for development of the model, camera filters are assumed to be narrow-band pass filters; i.e., the sensitivity is restricted to the central frequency of the broad-band pass filters (= λ_{cb}, where *cb* is R, G, or B):

$$S_{cb}(\lambda) = \begin{cases} 1 & if \ \lambda = \lambda_{cb} \\ 0 & if \ \lambda \neq \lambda_{cb}, \end{cases} \tag{3}$$

Assuming the camera to contain narrow-band pass filters, Eq. 2 reduces to

$$I_{cb} = I_{0,cb} \cdot e^{-A \cdot c_{cb}} \tag{4}$$

where $I_{0,cb}$ is the intensity of channel *cb* when no stain is present and c_{cb} is the absorption coefficient c_λ for $\lambda = \lambda_{cb}$.

The OD for a channel can be defined as

$$D_{cb} = -\ln\left(\frac{I_{cb}}{I_{0,cb}}\right) = A \cdot c_{cb}, \tag{5}$$

The OD for a channel D_{cb} depends linearly on the amount of stain, given the absorption value of the stain at channel *cb*. Usually, the overall intensity of the RGB signal is defined as

$$I = \frac{I_R + I_G + I_B}{3}. \quad (6)$$

Analogously, an overall measure for the OD can be defined as

$$D = \frac{D_R + D_G + D_B}{3} = \frac{A \cdot (c_R + c_G + c_B)}{3}. \quad (7)$$

Colour Models

RGB model. In the RGB model, the intensities I_R, I_G, and I_B, obtained by the camera for each pixel, are used. The combined I_R, I_G, I_B define a point in a Euclidean 3D coordinate space [RGB space (6); Appendix A]. Because of the limited dynamic range of the camera, the RGB space is limited to a cube (RGB cube). The perceived intensities depend on the absorption characteristics and the amount of a stain in a non-linear way (see Eq. 2). Therefore, different specimen locations containing a single stain in varying amounts will result in a non-linear relationship between the measured intensities I_R, I_G, and I_B for those locations. Intensities from pixels containing a single stain in varying amounts result in a complexly shaped 3D domain within the RGB cube (9,10). Stain recognition is possible by classifying each pixel to the stain corresponding with the 3D domain in which the RGB point is located. If the absorption characteristics of two stains are sufficiently different within the three sensitivity bands of the camera, differentiation based on the 3D RGB data will be possible.

In general, mathematical description of 3D domains requires complex, specialised and time-consuming algorithms. Therefore, mostly 1D or 2D projections of the 3D data are used for image segmentation (11). For example, handling (I_R, I_G) as coordinates in a plane, a stain can be described by identifying the appropriate 2D domain. Segmentation by applying an intensity threshold to one of the three camera channels corresponds to monochromatic thresholding.

HSI model. In general, for perceived light we can distinguish between intensity and chromaticity. Intensity depends on the amount of energy and is independent of the colour of the light. Chromaticity can be expressed as hue (dominant wavelength of the light, e.g., red, purple, etc.) and saturation (purity of the colour, e.g., pink is red with a low degree of saturation) (6). Using the equations from Appendix A, we can calculate values for hue, saturation, and intensity (HSI) from the RGB values acquired by the camera. This is achieved by projection of the 3D RGB vector onto a plane, so that the intensity component (as defined in Eq. 6) is split off. The resulting *chromaticity plane* $c_x c_y$ can be used to calculate values for H and S (Appendix A). In the chromaticity plane the RGB cube is represented by an equilateral triangle, which limits the extend of the $c_x c_y$ coordinates. As the effect of staining density variations on the HSI data is more easily understood in the $c_x c_y$ plane, this plane will be used throughout this study instead of the commonly used H and S. The 3D domains in the RGB space, resulting from a stain in varying amounts, will appear as a domain in the $c_x c_y$ plane. Classification of pixels as originating from a certain stain is possible by identifying in which of the domains, representing different stains, a pixel is located.

HSD model. For a single pixel, the values of D_{cb} (see Eq. 5) for the three channels result from the same amount of stain A. Therefore, for this pixel D_{cb} can be used as a relative measure for c_{cb}. The three D_{cb} span a new coordinate space $D_R D_G D_B$, which has the property that pixels resulting from one stain with varying amount will show a linear relationship. We can decouple the chromatic information in this space in the same way as we did in the $I_R I_G I_B$ space using the RGB to HSI transform (Appendix B). The component that has been split off now is the overall OD (see Eq. 7) instead of the intensity. Therefore, the newly derived model will be called the hue-saturation-density (HSD) model. The resulting chromatic coordinates $c_x c_y$ fully depend on the absorption coefficients c_R, c_G, and c_B. Again, the $c_x c_y$ coordinates are limited to the equilateral chromaticity triangle, caused by projection of the RGB cube.

Evaluation of Colour Models Using Simulated Data

Computer simulations were used to calculate theoretical values for I_R, I_G, and I_B and for c_x and c_y from both the HSI and HSD transforms for three different stains. Thus, the effect of (1) variations in the amount of stain, (2) mixtures of stains, and (3) broad-band instead of narrow-band camera filters could be studied in the three colour spaces in a controlled manner.

Using Eq. 2, expected intensities I_R, I_G, and I_B for three widely used stains were calculated. Calculations were performed both for the individual stains and for mixtures of stains. Computer simulations were used to approximate $c(\lambda)$ and $S(\lambda)$; $I_0(\lambda)$ was considered constant. Because only printed absorption curves were available for the Fast Red, diaminobenzidine tetrahydrochloride (DAB), and hematoxylin stains, values for $c(\lambda)$ were approximated by cubic spline interpolation through manually sampled points from the curves. The three approximated absorption curves are given in Figure 1.

For $S(\lambda)$, approximations of both narrow-band and broad-band pass filters were studied. The narrow-band pass filter was implemented as defined in Eq. 3, with λ_R = 600 nm, λ_G = 535 nm, and λ_B = 465 nm. Broad-band pass filters were approximated by Gaussian distributions, as shown in Figure 1, with $S_R(\lambda) = N(\mu = 600, \sigma = 23)$, $S_G(\lambda) = N(\mu = 535, \sigma = 24)$, and $S_B(\lambda) = N(\mu = 465, \sigma = 9)$ for the red, green, and blue channels, respectively.

Usually, the 10-based logarithm is used for calculation of the OD. OD values typically range from 0 to 2.5. In our case, in Eq. 5 the natural logarithm is used, and $c(\lambda)$ is arbitrarily scaled with maximum value smaller than 1. Therefore, in the computer simulations the amount A will range linearly from 0 to 5.75 (= 2.5 * ln(10)), corresponding to density 2.5 when the 10-based logarithm is used. Model calculations were performed using the Delphi programming environment (Borland, Scotts Valley, CA).

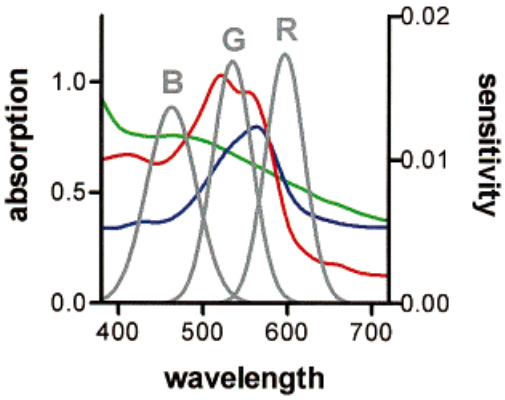

FIG. 1. Coloured lines show the absorption curves for three stains: Fast Red (red line), DAB (green line), and hematoxylin (blue line) (left y-axis); grey lines show the sensitivity of the three individual camera channels (red, green, and blue) of a three-chip CCD RGB camera (right y-axis).

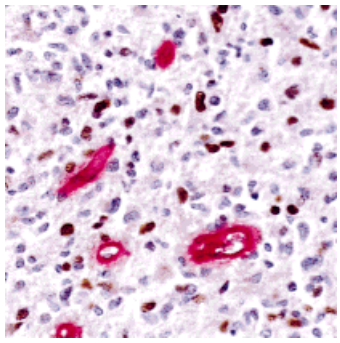

FIG. 2. Example of a microscopic image as used to evaluate the colour models, containing red-stained blood vessels, brown-stained nuclei of proliferating cells, and blue-stained nuclei of non-proliferating cells. Proliferating nuclei expressing little Ki-67 are weakly stained by the MIB1 antibody and also exhibit blue, non-specific staining of hematoxylin.

Graphical representations were produced using Graphpad Prism 2.01 (Graphpad Software, Inc., San Diego, CA).

Evaluation Using Experimental Data

Immunohistochemistry. Immuno-double staining techniques are used for the simultaneous detection of two antigens. In these techniques, DAB and Fast Red are often used to visualise the two antigens, whereas hematoxylin is used for counterstaining. An example of immuno-double staining of a tissue section is shown in Figure 2. In this example, immuno-double staining was used for the simultaneous detection of blood vessels and proliferating cells in a 4 μm thick, formalin-fixed and paraffin-embedded tissue section of a biopsy specimen of a supratentorial diffuse astrocytic neoplasm. Antibodies that are directed to collagen IV and Ki-67 (MIB1) were used to visualise the blood vessels and proliferating cells, respectively. The chromogenes Fast Red and DAB were used to stain the blood vessels and proliferating cells, respectively, and hematoxylin was used as counterstain. Details of the immunohistochemical technique are given elsewhere (12).

Colour image analysis in immunohistochemically stained tissue sections. To study the RGB, HSI, and HSD colour domains of the stains in the immunohistochemically stained tissue sections, images were recorded by a three-chip CCD RGB camera (DXC-325P, Sony) mounted on a conventional light microscope (Axioskop, Carl Zeiss, Weesp, The Netherlands), using a ×10 objective (NA = 0.3, resulting pixel size 0.78 × 0.83 μm). Spectral characteristics of the colour filters in our camera were not available. Image acquisition and analysis were performed using a Vidas^{plus} system (Kontron GMBH, Eching, Germany). Köhler illumination was used and a warm-up time for camera and microscope of at least 100 min (13) was applied. Manually selected microscopic fields were digitised and stored on magneto-optical disc (Borsu Systema, Lelystad, The Netherlands) as true colour (24-bit RGB) images. Nuclei of non-proliferating cells, nuclei of proliferating cells, and immunohistochemically stained blood vessels were manually selected in the stored images with a computer mouse. Thus, RGB values for hematoxylin, DAB, and Fast Red were obtained for locations originating from varying amounts of stain.

Because chromatic information for pixels with low staining density (i.e., high transmitted intensity) is intrinsically unstable (14), a minimum OD threshold of 0.15 was applied before scatter plots of RGB, HSI, and HSD data were constructed. Prior to the recording of images of tissue specimens, an RGB image of an empty microscopic field and an RGB dark-current image were stored. For the RGB and HSI models, the R, G, and B images were shading corrected using the stored empty field and dark-current images, according to an algorithm described elsewhere (15). For the HSD model, the stored image of an empty field was used for determining the amount of light entering the specimen for each pixel ($I_{0,cb}$). Thus, correction for unequal illumination is implicitly performed. Correction for dark current was applied in the HSD model by subtracting the pixel values of the dark-current reference image from the pixel values of I_{cb} and $I_{0,cb}$ for all three camera channels.

RESULTS

Simulated Data

RGB model. Combined red and green intensities ($I_R I_G$ plane) for the simulated Fast Red, DAB, and hematoxylin stains are shown in Figure 3a. In a previous study (16), it was found that this particular 2D projection of the 3D RGB data was adequate for differentiation between Fast Red and hematoxylin. Data for both the narrow-band and broad-band camera models are shown.

Points originating from a single stain with varying amounts are located on a curve, displaying the non-linear relationship between the amount of stain and the perceived light intensity. Pixels originating from DAB and hematoxylin can be discerned from pixels originating from Fast Red, but cannot be discerned from each other (Fig. 3a). Dashed lines indicate the location of mixtures of hematoxylin and Fast Red for two different amounts of

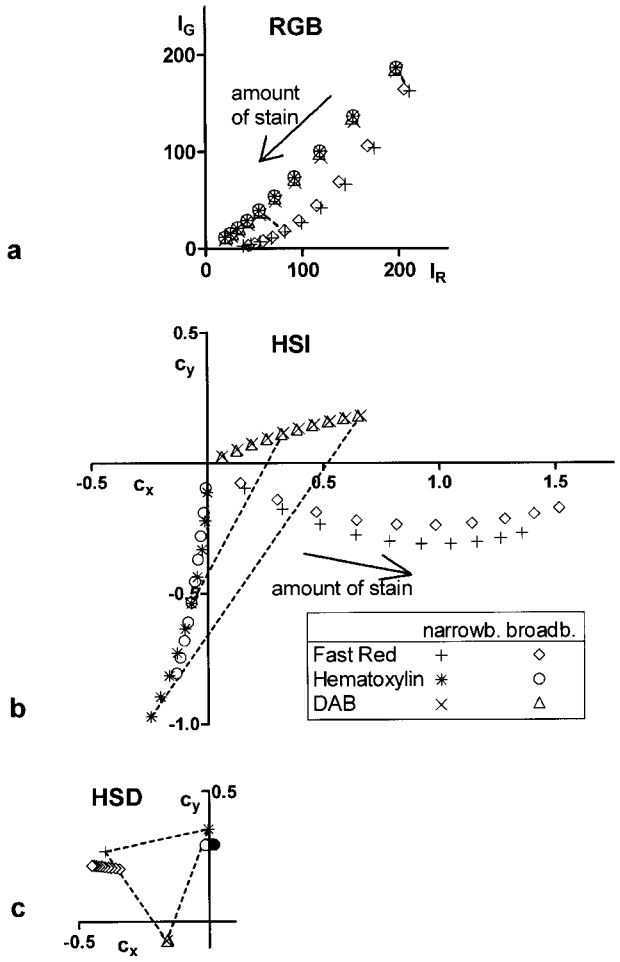

FIG. 3. **a:** The theoretical red and green intensities of the Fast Red, hematoxylin, and DAB stains in the test image, assuming narrow-band pass (represented by plus, asterisk, and cross symbols, respectively) or broad-band pass (diamond, circle, and triangle symbols, respectively) camera filters. An arrow indicates the direction of increasing stain density, linearly ranging from 0.575 to 5.75. Dashed lines show the theoretical effect of mixing hematoxylin and Fast Red for joint stain density 2.875 and 5.75. **b:** The theoretical c_xc_y data of the HSI transform applied to the intensities of a. Dashed lines show the theoretical effect of mixing DAB and hematoxylin for joint stain density 2.875 and 5.75. **c:** The theoretical c_xc_y data of the HSD transform applied to the intensities of a. Dashed lines show the theoretical effect of all possible mixtures of two stains.

stain. In general, pixels originating from a mixture of two stains result in trajectories interconnecting the curves of the individual stains. The location of a point on this trajectory is determined by the relative amount of individual stains in the mixture. The total amount of stain in a mixture determines the location of such a trajectory. Comparison between the narrow-band and broad-band camera models shows only minor differences.

HSI model. The result of applying the RGB to HSI transform to the theoretical intensities is shown in Figure 3b. Pixels originating from the three stains are located on curves, which extend from the origin. For all three stains, chromatic coordinates c_xc_y from the HSI transform are highly dependent on the amount of stain. Curves observed for the three pure stains are distinct from each other and thus the stains can be discerned.

Dashed lines show the theoretical trajectories resulting from a mixture of DAB and hematoxylin for two different amounts of stain. As was seen for the RG plane, the amount of stain in the mixture determines the location of the trajectory, whereas the relative amounts of the individual stains in the mixture determine the location of a point on this trajectory. In this way, pixels originating from a mixture of DAB and hematoxylin with varying amounts of the individual stains result in a large 2D domain in the c_xc_y plane. This domain partly overlaps the Fast Red curve. Therefore, unequivocal classification of pixels using the HSI model is not always possible when these particular stains are used. Comparison of the results obtained by narrow-band and broad-band filters discloses different effects for the three dyes. No notable effect is present for DAB, whereas the effects are most prominent for Fast Red.

HSD model. Theoretical RGB intensities from a stain in varying amounts result in a single point in the c_xc_y chromaticity plane of the HSD transform (Fig. 3c), assuming narrow-band pass filters in the camera. This shows that the c_xc_y plane of the HSD transform only contains information regarding the spectral characteristics of the stains. Discrimination between stains is possible if the absorption curves are sufficiently different. Dashed lines show results of mixtures of all possible combinations of two stains for the narrow-band camera filter model. Pixels originating from a mixture of two stains will be located on a line connecting the points resulting from the individual stains. The location of such a pixel is only affected by the relative amount of each of the stains, and does not depend on the total amount of stain. Using these three respective stains, unequivocal classification of pixels as belonging to stains or mixtures of stains is possible. This is shown mathematically in Appendix C.

The effect of broad-band camera filters is stain dependent; for the hematoxylin stain hardly any effect is seen, whereas for the Fast Red stain the largest effect is visible. This effect comprises a small dependence between chromatic coordinates and the amount of stain and shows a shift from the data cluster to a slightly different location. The effect of applying broad-band pass filters in the camera is smaller for the HSD model than it is for the HSI model.

Experimental Data

RGB model. RGB data resulting from measurement in an immuno-doublestained section are given in Figure 4a,b. Pixels originating from nuclei of non-proliferating cells stained by hematoxylin, from nuclei of proliferating cells stained by DAB, and from blood vessels stained by Fast Red are clustered in elongated domains in the I_RI_G plane (Fig. 4a). The domains for DAB and especially the one for Fast Red show a curved shape. This resembles the curved patterns seen for computer simulations (Fig. 3a). Because

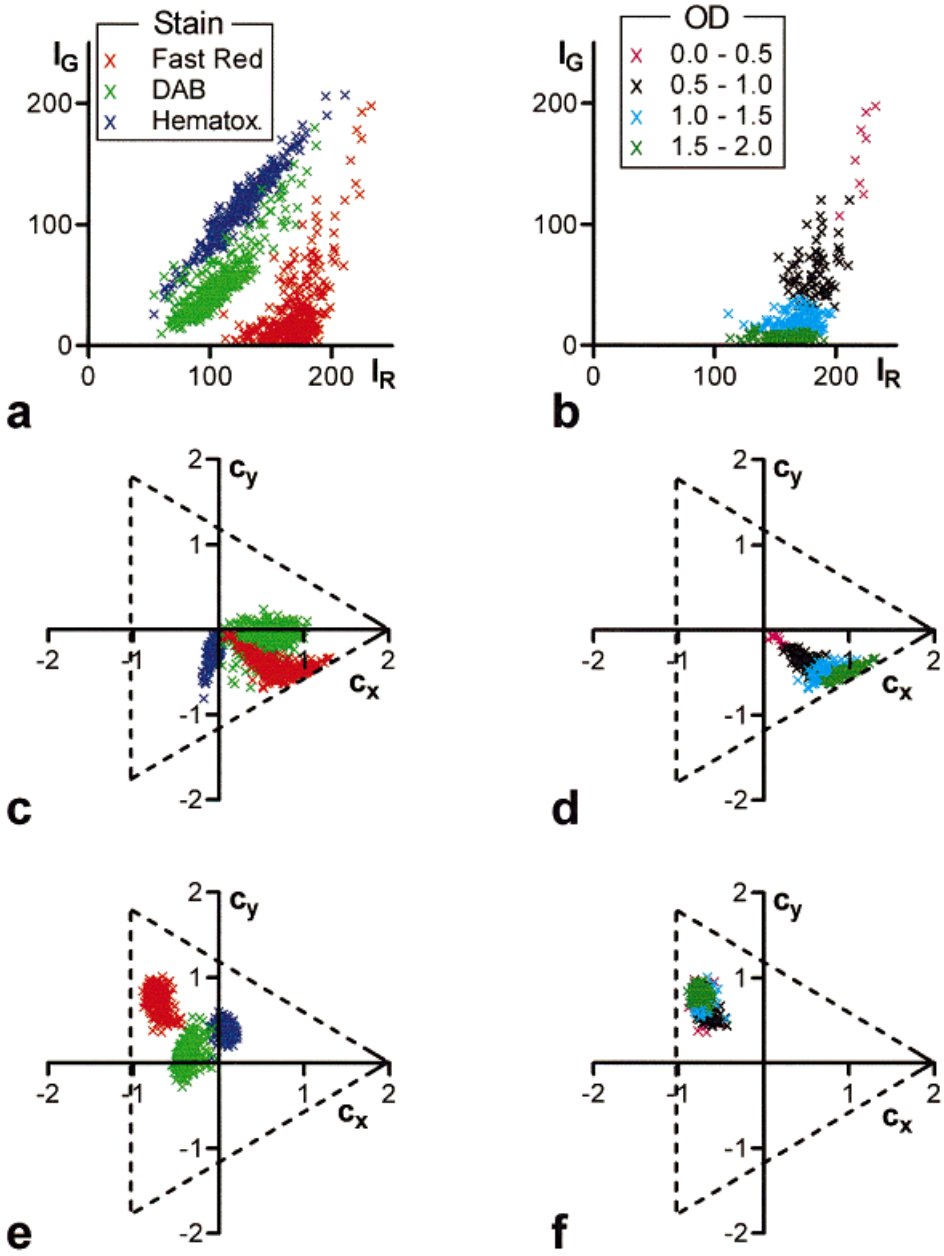

FIG. 4. **a:** The 2D projection of the 3D RGB data of the three stains in the test image onto the red-green plane. **b:** The 2D projection of RGB data of the Fast Red staining, divided into four OD classes, revealing that the localisation within the Fast Red domain depends on the OD. **c:** The chromatic $c_x c_y$ coordinates of the three stains in the test image after the HSI transform. The dashed line shows the triangular restriction of the plane (see Appendix A). Note the considerable overlap between the red and green pixels. **d:** Subdivision of the HSI data of the Fast Red stain into four OD classes. **e:** The chromatic $c_x c_y$ coordinates of the three stains in the test image after the HSD transform are localised in well-defined domains that exhibit little overlap. **f:** Subdivision of the HSD data of the Fast Red stain into four OD classes, showing that the localisation within these domains is less dependent on the OD of the stain.

of system noise and variations in the spectral characteristics of the stains, the single lines from Figure 3a now show as domains. The DAB and Fast Red domains exhibit hardly any overlap, whereas the DAB and hematoxylin domains do show some overlap. The position of pixels within these domains depends on the OD and consequently on the amount of stain, as shown for Fast Red in Figure 4b.

Comparing results of computer simulations with experimental data, we notice a complete overlap between hematoxylin and DAB for theoretical data (Fig. 3a), whereas these stains are almost entirely separable in the experimental data (Fig. 4a). This difference may be attributed to two factors: (1) the spectral characteristics of the filters in our CCD camera were not known, and were therefore approximated with Gaussian distributions in the computer simulations; or (2) the absorption curves for the three stains we used were obtained from measurement in solution, and may change after application to a tissue section.

HSI model. Applying the HSI transform to the RGB intensities from the three stains in the immunohistochemically stained specimen results in curved domains in the chromaticity triangle, representing the individual stains (Fig. 4c). Considerable overlap can be seen between the Fast Red domain and the DAB domain. Because in the RGB space we can easily distinguish between these two stains, we conclude that in the $c_x c_y$ coordinates from the HSI transform not all chromatic information, present in the RGB data, is retained. Use of the HSI model does not enable all possible distinctions between stains in a 2D coordinate space.

Subdividing the Fast Red domain in OD classes (Fig. 4d) shows that the chromatic coordinates $c_x c_y$ are strongly influenced by the amount of stain. For low OD, domains of all three stains show a trend toward the origin of the chromaticity triangle, showing that discrimination between background pixels and pixels originating from a small amount of stain will be troublesome in the HSI model.

HSD model. Pixels originating from the three stains in the immunohistochemically stained specimen result in well-demarcated colour domains in the $c_x c_y$ plane of the HSD transform (Fig. 4e). Hardly any overlap can be seen between these clusters, showing that the separability between stains in the RGB data is retained in the chromatic coordinates of the HSD model. The position of a point within a colour domain is less influenced by the amount of stain bound to the specimen than was observed for the RGB or HSI model (Fig. 4f). Because in Figure 4f points originating from small amounts of Fast Red (small OD) do not show a trend toward the origin (which represents parts of the image where no stain is present, i.e., the background), even pixels originating from small amounts of stain will be classified correctly.

DISCUSSION

For application of image analysis techniques in transmitted light microscopy, reliable segmentation of the objects of interest is of major importance (7). The first step in recognition of stained objects is classification of pixels as originating from one of the used stains (17). When only one stain is present in an otherwise clear background, pixel classification may be performed using monochromatic thresholding (7). Using immunohistochemical techniques results in general in specimens containing at least two different stains. Monochromatic thresholding will not suffice for this distinction. In previous studies, a number of approaches have been applied in which stain recognition was performed using the OD measured simultaneously at two or more different wavelengths, either separately (18,19) or combined (1,20). The disadvantage of such methods is that expensive special hardware is needed (suitable filters, automatic filter wheel) (21) and that results may be suboptimal in case of slight staining variations (22,23). Use of an RGB colour camera is an attractive alternative because this device is generally applicable and widely available. Because of its analogy with the human eye, images acquired by such a device are suitable for visual inspection by humans (23), possibly enabling interactive correction of images.

For stain recognition in RGB colour images, algorithms are needed to perform pixel classification in the RGB space. To avoid the use of complicated algorithms handling 3D data, in most previous studies stain recognition is performed in the RGB model either by monochromatic thresholding in separate channels (5,24) or in a projection of the 3D data onto a plane (16) or a line (7,21,23). As was shown in the present study, use of the RGB model will only result in solutions that are generally applicable if the entire 3D data space is used. Projections onto a line or plane can be used for specific applications but will not be generally applicable. In addition, staining variations may hamper results. In a number of studies, the RGB to HSI transform is performed to be able to perform stain recognition in a standardised 2D data space (2,17,25–28). As was shown, the HSI model is not adequate for application in transmitted light microscopy, because in this model chromatic information is acquired by decoupling light intensity instead of OD.

An algorithm was described by Garbay et al. (29) in which the HSI transform was applied to log-transformed intensities acquired by a colour RGB camera. Log-transformed intensities were used in order to remain close to the way the human eye processes colour. In our opinion, a colour model used for recognition of stains in transmitted light microscopy should not be based on the human visual system. Instead, it should provide chromatic information independent of the OD of the stain. Therefore, in the HSD model, application of the RGB to HSI transform to ODs was chosen in analogy with Lambert-Beers' law of absorption.

Because all available information concerning an image from a three-chip CCD RGB camera is present in the RGB data, use of alternative colour spaces (either well-known standardised spaces like the HSI space or custom projections) is only useful to increase availability of the desired information by removing information that hampers this distinction. RGB data contain a mixture of chromatic (spectral characteristics of used stains) and achromatic (amount of stain, expressed in the OD) information (25). Moreover, this information is mixed in a non-linear way. For classification of pixels as belonging to a certain stain, the only information useful is the information that differentiates the specific stain from other stains or from the background, being the absorption characteristics of the stain. Ideally, other factors should be removed entirely to facilitate correct pixel classification. Acquisition of a microscopic image by a CCD colour camera can be regarded as sampling of the absorption curve over three predefined regions in the visible light spectrum. Because the absorption curves are arbitrarily scaled, these three samples can effectively be reduced to two numbers by normalisation (e.g., by scaling the sum of the three samples to a constant). The HSD model, introduced in the present study, enables pixel classification in a standardised 2D data space, which contains all spectral information independent of the amount of stain. Thus, the HSD model fulfills

the above formulated requirement. A substantial improvement may be expected when the complete spectral information of a stain can be used (23). Recently, a device has been developed which enables acquisition of high-resolution spectral information (30). Using sophisticated algorithms to perform stain recognition, results using such a device may be more accurate than those using a CCD colour camera, especially when many stains are present in a single image. Because it is widely available and easy to use, the CCD RGB camera remains an interesting alternative.

Theoretically, after decoupling the OD, the $c_x c_y$ plane of the HSD transform contains all spectral information, enabling all possible distinctions between stains based on spectral information only. The model simulations illustrate this, where (assuming narrow-band camera filters) each stain is present as a single dot. Mixtures of stains show linear trajectories between the single dots. The $c_x c_y$ coordinates do not depend on the amount of stain, both for mixtures and for single stains. Use of broad-band filters was found to cause a small deviation from the theoretical single dot in the HSD model. Data from the experimental study confirm the superiority of the HSD model, because each stain is represented by a well-defined cluster of points in the $c_x c_y$ plane of the HSD transform. Only a small amount of overlap was observed between the hematoxylin cluster and the DAB cluster, which was present in the $I_R I_G$ plane as well. Using experimental data, a small dependence was observed between chromatic coordinates and the OD of the stain. This effect may be attributed to two factors: (1) the fact that the RGB values from locations containing very high or extremely low amounts of stain are not accurate; or (2) the presence of broad-band filters in the camera, as shown in the model simulations.

In this article, an adaptation of the well-known RGB to HSI transform is proposed, which consists of decoupling the OD instead of the light intensity from the RGB data. This newly derived colour transformation, named the RGB to HSD transform, was found to enable the recognition of light-absorbing stains in transmitted light microscopic images in a 2D data space in a standardised manner. Thus, use of complicated algorithms dealing with 3D data is avoided, while keeping the full ability to differentiate between stains that are sufficiently different in the RGB space.

ACKNOWLEDGMENTS

The authors express their gratitude to the personnel from the Laboratory for Analytical Cellular Pathology for their excellent immunohistochemical stainings. Also, the authors thank Dr. M. Macville for his valuable contribution.

LITERATURE CITED

1. Tucker J, Schenck U, Eiermann W, Burger G. Quantitative image analysis on immunocytochemically labeled cells. In: Burger G, Ploem JS, Goerttler K, editors. Clinical cytometry and histometry. San Diego: Academic Press; 1987. p 402–406.
2. Fermin CD, Gerber MA, Torre-Bueno JR. Colour thresholding and objective quantification in bioimaging. J Microsc 1992;167:85–95.
3. Wied GL, Bartels PH, Bibbo M, Dytch HE. Image analysis in quantitative cytopathology and histopathology. Hum Pathol 1989;20:549–571.
4. Furness PN. The use of digital images in pathology. J Pathol 1997; 183:253–263.
5. Rostagno P, Birtwisle I, Ettore F, Courdi A, Gioanni J, Namer M, Caldani C. Immunohistochemical determination of nuclear antigens by colour image analysis: application for labelling index, estrogen and progesterone receptor status in breast cancer. Anal Cell Pathol 1994; 7:275–287.
6. Gonzalez RC, Woods RE. Digital image processing. Reading, MA: Addison-Wesley; 1992.
7. MacAulay C, Tezcan H, Palcic B. Adaptive color basis transformation: an aid in image segmentation. Anal Quant Cytol Histol 1989;11:53–58.
8. Longhurst RS. Geometrical and physical optics. London: Longmans; 1964.
9. Sarabi A, Aggarwal JK. Segmentation of chromatic images. Pattern Recognition 1981;13:417–427.
10. Acosta G, Waks A. Multi-spectral acquisition and analysis of FISH signals. In: Benedetti PA, Celli RM, Evangelista V, Guidarini D, Vestri S, editors. Microscope instrumentation. CA-AMCA E.C. Concerted action: automation in molecular cytogenetic analysis. Pisa: 1994. p 9–15.
11. Celenk M. A color clustering technique for image segmentation. Comp Vision Graph Im Proc 1990;52:145–170.
12. Wesseling P, van der Laak JAWM, Link M, Teepen HLJM, Ruiter DJ. Quantitative analysis of microvascular changes in diffuse astrocytic neoplasms with increasing grade of malignancy. Hum Pathol 1998; 29:352–358.
13. Willemse F, Nap M, de Kok LB, Eggink HF. Image analysis in immunohistochemistry: factors with a possible influence on the performance of Vidas version 2.0, a commercially available true color image analysis system. Anal Quant Cytol Histol 1993;15:136–143.
14. Perez F, Koch C. Towards color image segmentation in analog VLSI: algorithm and hardware. Int J Comp Vision 1994;12:17–42.
15. de Meester U, Young IT, Lindeman J, van der Linden HC. Towards a quantitative grading of bladder tumors. Cytometry 1991;12:602–613.
16. van der Laak JAWM, Westphal JR, Schalkwijk LJM, Pahlplatz MMM, Ruiter DJ, de Waal RMW, de Wilde PCM. An improved procedure to quantify tumour vascularity using true colour image analysis. Comparison with the manual hot-spot procedure in a human melanoma xenograft model. J Pathol 1998;184:136–143.
17. Poston RN, Gall NP. Hue-saturation-intensity colour image analysis for the quantitation of immunoperoxidase staining. In: Elder HY, editor. Micro 90. Bristol: Adam Hilger; 1990. p 525–528.
18. Figge J, Bakst G, Weisheit D, Solis O, Ross JS. Image analysis quantitation of immunoreactive retinoblastoma protein in human thyroid neoplasms with a streptavidin-biotin-peroxidase staining technique. Am J Pathol 1991;139:1213–1219.
19. Weinberg DS. Quantitative immunocytochemistry in pathology. In: Marchevsky A, Bartels PH, editors. Image analysis: a primer for pathologists. New York: Raven Press; 1994. p 235–260.
20. Aggarwal RK, Bacus JW. A multi-spectral approach for scene analysis of cervical cytology smears. J Histochem Cytochem 1977;25:668–680.
21. Ruifrok A. Quantification of immunohistochemical staining by color translation and automated thresholding. Anal Quant Cytol Histol 1997;19:107–113.
22. Brugal G. Image analysis of microscopic preparations. Meth Achiev Exp Pathol 1984;11:1–33.
23. Harms H, Aus HM, Haucke M, Gunzer U. Segmentation of stained blood cell images measured at high scanning density with high magnification and high numerical aperture optics. Cytometry 1986; 7:522–531.
24. Bulten J, van der Laak JAWM, Gemmink JH, Pahlplatz MMM, de Wilde PCM, Hanselaar AGJM. MIB1, a promising marker for the classification of cervical intraepithelial neoplasia. J Pathol 1996;178:268–273.
25. Lamaziere JMD, Lavallee J, Zunino C, Larrue J. Semiquantitative study of the distribution of two cellular antigens by computer-directed color analysis. Lab Invest 1993;68:248–252.
26. Barbareschi M, Girlando S, Mauri FM, Forti S, Eccher C, Mauri FA, Togni R, Dalla Palma P, Doglioni C. Quantitative growth fraction evaluation with MIB1 and Ki67 antibodies in breast carcinomas. Am J Clin Pathol 1994;102:171–175.
27. Weaver JR, Au JLS. Application of automatic thresholding in image analysis scoring of cells in human solid tumors labeled for proliferation markers. Cytometry 1997;29:128–135.
28. Kohlberger PD, Obermair A, Sliutz G, Heinzl H, Koelbl H, Breitenecker G, Gitsch G, Kainz C. Quantitative immunohistochemistry of

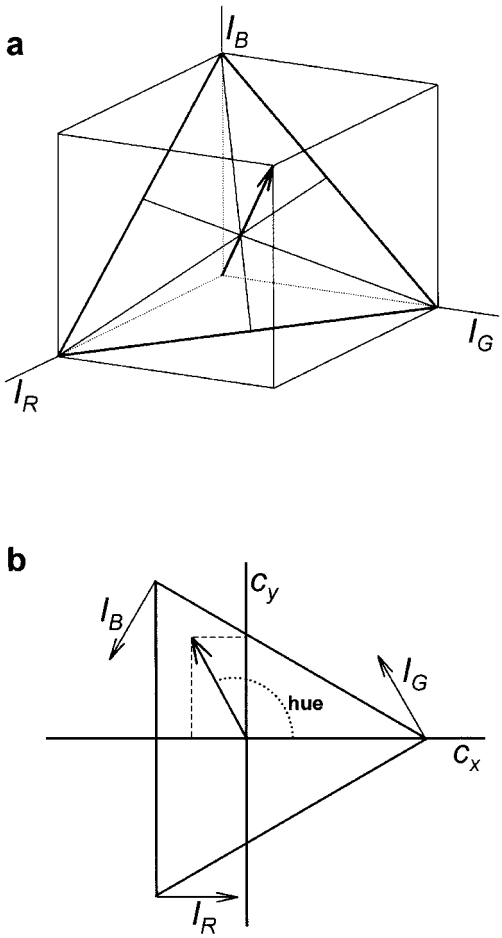

FIG. 5. **a:** The RGB camera signal can be viewed as a 3D coordinate space. Camera limitations limit the space to a cube (RGB cube); also shown is the triangle with constant overall intensity equal to the maximum of the individual channel intensities. **b:** The plane resulting from the projection described in Appendix A, with indicated new coordinate system $c_x c_y$.

factor VIII-related antigen in breast carcinoma: a comparison of computer-assisted image analysis with established counting methods. Am J Clin Pathol 1996;105:705–710.
29. Garbay C, Brugal G, Choquet C. Application of colored image analysis to bone marrow cell recognition. Anal Quant Cytol 1981;3:272–280.
30. Garini Y, Macville M, du Manoir S, Buckwald RA, Lavi M, Katzir N, Wine D, Bar-Am I, Schrock E, Cabib D, Ried T. Spectral karyotyping. Bioimaging 1996;4:65–72.

APPENDIX A
Formal Definition of the HSI Model

In analogy with the human eye, a CCD colour camera splits a colour into its composing primaries red, green, and blue. These three primary components can be depicted in a 3D coordinate space, in which the axes represent the intensities of the primary colours. Because of the limited dynamic range of the camera, the intensity domain is limited to a cube in the RGB space (Fig. 5a). The origin of the RGB space represents black (i.e., all three primary intensities are zero), and the farthest corner of the cube represents white (all primary intensities maximal). The line connecting those extremes represents all grey values. Each composite colour is represented by a point in this cube. The intensity of a composite colour is defined as $I = (I_R + I_G + I_B)/3$, the average of the three primary intensities. In the RGB cube, each plane perpendicular to the grey-scale diagonal has the property that points located on the plane have equal intensity (Fig. 5a shows one such plane). Points on such a plane are limited to a triangular domain, of which the size linearly depends on the value of I for the particular plane. The grey-scale diagonal intersects this triangle in the centre of gravity. It is possible to normalise this equilateral *chromaticity triangle* to a standardised size, making the coordinates of a point on this plane independent of the intensity.

In the chromaticity triangle, a 2D coordinate system $c_x c_y$ can be defined with the centre of gravity as origin and with the positive x-axes through the corner where the red intensity is maximal (Fig. 5b). Every point in the RGB space can be transformed to its chromaticity coordinates, which are independent of the intensity. For every two RGB points having the same chromaticity coordinates, the ratios between the RGB intensities are identical. Calculation of the chromatic coordinates requires projection of the RGB data onto the triangle. This can be performed using the following equations (derivations not given):

$$c_x = \frac{I_R}{I} - 1$$

$$c_y = \frac{I_G - I_B}{\sqrt{3} \cdot I}.$$

These coordinates can be viewed as a vector from the origin to a point on the plane. The *saturation* is defined as the length of this vector, and the *hue* as the angle of the vector with the positive x-axis:

$$hue = \arctan\left(\frac{c_y}{c_x}\right)$$

$$saturation = \sqrt{c_x^2 + c_y^2}.$$

Back transformation from HSI to $c_x c_y$ may be performed using the equations:

$$c_x = saturation \cdot \cos(hue)$$

$$c_y = saturation \cdot \sin(hue).$$

Back transformation from $c_x c_y$ to RGB may be performed using the equations:

$$I_R = I \cdot (c_x + 1)$$

$$I_G = \frac{1}{2} \cdot I \cdot (2 - c_x + \sqrt{3} \cdot c_y)$$

$$I_B = \frac{1}{2}\cdot I\cdot(2-c_x-\sqrt{3}\cdot c_y).$$

APPENDIX B
Formal Definition of the HSD Transform

By using the projection from Appendix A on the ODs of the three channels (see Eq. 5) instead of the intensities, the RGB to HSD transform is defined:

$$c_x = \frac{D_R}{D} - 1$$

$$c_y = \frac{D_G - D_B}{\sqrt{3}\cdot D}.$$

Note that because the OD is decoupled rather than the intensity, the chromatic coordinates of the HSD model are not equal to those of the HSI model. For the HSD model, the resulting $c_x c_y$ plane has the property that single points result from RGB points with identical ratios between the c_R, c_G, and c_B. Thus, all information regarding the absorption curves is represented in a single plane. In analogy with the HSI model, values for hue and saturation can be calculated from the chromaticity triangle. Because mixtures of stains show a linear pattern in the $c_x c_y$ plane of the HSD model (Appendix C), it is recommended to use this plane instead of hue and saturation.

Back transformation from $c_x c_y$ to individual channel densities D_R, D_G, and D_B may be performed using the equations from Appendix A. Individual channel densities can be transformed back to RGB using the equation:

$$I_{cb} = I_{0,cb}\cdot e^{-D_{cb}}.$$

If the empty reference image $I_{0,cb}$ (for cb = R, G, and B) is available, the original RGB values can be calculated.

APPENDIX C
Mixtures of Stains in the HSD Model

Assume two pure stains, with given amount A_i and absorption characteristic $c_{i,cb}$ for stain i = 1,2; camera channel cb = R, G, or B. Then, according to Eq. 4, for the narrow-band model, a mixture of the stains will result in camera intensity:

$$I_{M,cb} = I_{0,cb}\cdot e^{-A_1\cdot c_{1,cb}}\cdot e^{-A_2\cdot c_{2,cb}} = I_{0,cb}\cdot e^{-(A_1\cdot c_{1,cb}+A_2\cdot c_{2,cb})}$$

where $I_{M,cb}$ is the intensity of camera channel cb for the mixture of two stains. The channel density (see Eq. 5) now becomes:

$$D_{M,cb} = -\ln\left(\frac{I_{M,cb}}{I_{0,cb}}\right) = A_1\cdot c_{1,cb} + A_2\cdot c_{2,cb} = D_{1,cb} + D_{2,cb}$$

with $D_{i,cb}$ the OD for the individual camera channel cb, for staining i (i = 1,2). The overall OD of the mixture, D_M, now becomes:

$$D_M = \sum_{i=1,2}\frac{D_{i,R} + D_{i,G} + D_{i,B}}{3} = D_1 + D_2$$

with D_i the individual overall density of staining i. Applying the vector projection from Appendix A to the mixed stains gives:

$$c_{M,x} = \frac{D_{M,R}}{D_M} - 1 = \frac{D_{1,R} - D_1 + D_{2,R} - D_2}{D_M}$$

$$= \frac{D_1}{D_M}\cdot c_{1,x} + \frac{D_2}{D_M}\cdot c_{2,x} = \frac{D_1}{D_M}\cdot c_{1,x} + \left(1 - \frac{D_1}{D_M}\right)\cdot c_{2,x}$$

with $c_{i,x}$ the c_x coordinate for the individual staining i. Also, for c_y:

$$c_{M,y} = \frac{D_{M,G} - D_{M,B}}{D_M} = \frac{D_1}{D_M}\cdot c_{1,y} + \frac{D_2}{D_M}\cdot c_{2,y}$$

$$= \frac{D_1}{D_M}\cdot c_{1,y} + \left(1 - \frac{D_1}{D_M}\right)\cdot c_{2,y}.$$

So, the points $(c_{M,x}, c_{M,y})$ are on the line segment connecting $(c_{1,x}, c_{1,y})$ with $(c_{2,x}, c_{2,y})$. The position of the point along the line is uniquely determined by D_2/D_1.