Accurate Facial Colour Determination and Segmentation Compensated for non-uniform Image Lighting, Geometry and Camera Characteristics.

Tony Adriaansen
CSIRO
Telecommunications & Industrial Physics
PO Box 76, Epping, NSW 2121, Australia
tony.adriaansen@tip.csiro.au

Abstract

In determining facial colour and segmentation, a novel method is presented which compensates for external factors affecting image colour measurements. The method also allows robust and accurate colour differences to be made between images, captured under differing conditions. It has been shown that the typical approach using normalised RGB for facial colour determination may be inadequate in accurately finding faces. This paper presents results on facial skin colour analysis of 33 subjects spanning three groups, resulting in a more accurate and robust segmentation than is possible using the approach typically used.

Group1: {Australian & Northern European}. Group2: {Sri Lankan, Indian & Egyptian}. Group3: {Chinese, Vietnamese & Taiwanese}.

1. Introduction

In medical image analysis it is often necessary to accurately locate and measure face colour. An object's colour in an image results from a non-linear combination of different absorption & scattering properties over the visible wavelength range, the spectral properties of the illuminant, the spatial geometry of the object with respect to the illuminant and detector, and the spectral response of the detector. A change in any of these parameters has the effect of modifying the object colour, and the consequent colour values stored in a captured digital image.

As is widely known, the Commission Internationale de l'Eclairage (CIE) set standards in 1931 for colour measurement, by which any colour may be characterised by a set of 3 functions known as the XYZ tristimulus matching functions, after specifying the parameters above. It is noted that all colour space coordinate scales can be derived from the XYZ tristimulus measurement system.

A common method used in attempting to compensate images due to changes in the above parameters is by introducing a (typically) white or grey patch into the field of view and linearly interpolating the image {red, green, blue} vector of the patch area to match an initial image. In fact, Hall et al. [1] found that it is not sufficient to simply have a reference white and black in the image for "calibration" purposes, as this would assume a linear relationship for all shades of grey in between. Such greyscale non-linearities are inherent in all imaging systems. Referencing to a white or grey patch does go some way towards normalising image intensity, however, further work is necessary to ensure the correct reproduction of colour, as well as intensity, Frey and Palus[2]. Others [3], [4], [5], [6] have considered the measurement of a colour in a digital imaging system and outlined a method for calibration. In particular, they state that greyscale linearisation of each of the three channels {R,G,B} is not enough to allow the system to reproduce colours or hues correctly.

This paper addresses the inherent limitations of the common approach described above where each of the three {R,G,B} channels are treated in isolation, and outlines prerequisites and a method which preserves the inherent tri-dimensionality of colour images, resulting in robust colour image measurements.

2. Method

2a. Overview

A reference system consisting of a light source, a CCD camera detector and a minimum of 3 reference colours located at specified positions relative to each other is defined. The system is then used as a reference from which a robust "truth set" can be generated. Positional geometry is important for colour measurement, mainly to avoid specular reflection. Such reflection offers no colour information (except in the special case of metals).

The colour measurement signal in an image is determined by the following equation:

$$P_i = \int_\lambda E_{i(\lambda)} . R_{i(\lambda)} . S_{(\lambda)} \, d\lambda$$

Equation (1)

where P_i = image value.
E = energy of illumination, light source.
R = spectral response of detector, CCD camera.
S = spectral reflectance of an object, face.
i = band filter response, camera R, G, B.
λ = wavelength, visible range.

ie: Pi depends on the 3 variables E, R and S.

In this paper, only light sources, that approximate the spectral response of a perfect blackbody radiator[8], are used.

In the design of a robust system there are two aspects which must be considered, the light source and the detector.

2b.1 Constraints: Light source.

Figure 1a shows the Planckian locus of a blackbody on a chromaticity diagram, for temperatures ranging from about 1700 Kelvin to infinity with points shown at 2850 and 6500 Kelvin. The sail shape defines the boundary of all colours without the luminance or brightness (CIE)Y component. Figure 1b shows corresponding dominant wavelengths, 400nm to 700nm

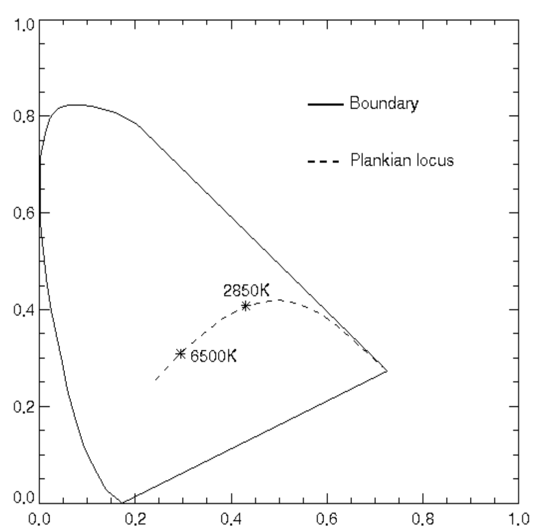

Figure 1a

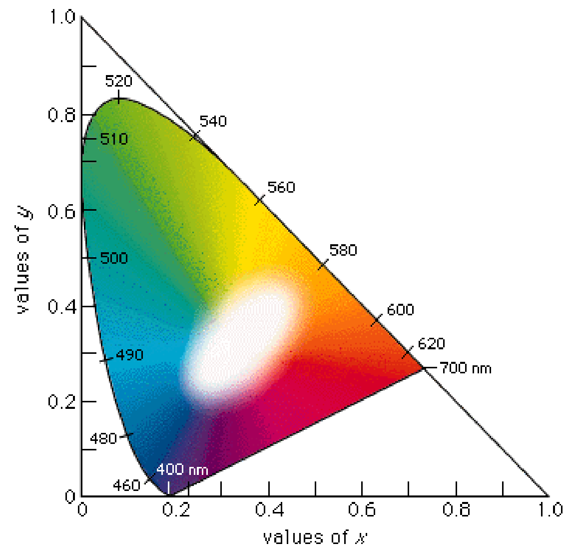

Figure 1b

The spectral response of a (blackbody-type) light source can be described by its “correlated colour temperature”; ie one where it’s spectral response is similar to a blackbody at a given temperature. It can be seen in Figure 1 that low colour temperature sources (eg. 1700 Kelvin) have a more reddish dominant wavelength whereas higher colour temperature sources tend more towards blue. The relationship between dominant wavelength λ(max) to temperature of a blackbody[8] is:

$$\lambda_{(max)} . T_{(Kelvin)} = 2891_{(micron\ Kelvin)}$$

Equation (2)

It is known that the blackbody spectral emissive power curves[8] have a highly positive skewness factor, sharply decreasing for wavelengths less than λ(max), therefore for colour measurements spanning the visible range of 380nm to 720nm, a light source having a high correlated colour temperature is desirable.

2b.2 Detector constraints.

Detector response Ri is a function of the Silicon CCD sensor, glass camera lens, filters and amplifier electronics. The camera employs a “standard front side illuminated CCD” with spectral response as shown in Figure 2a. Normalised lens spectral response percent transmission is shown in Figure 2b.

CCD Spectral Response Characteristics

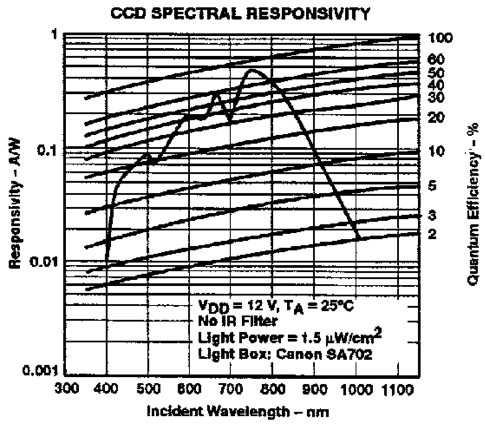

Figure 2a

Lens Spectral Response Characteristics

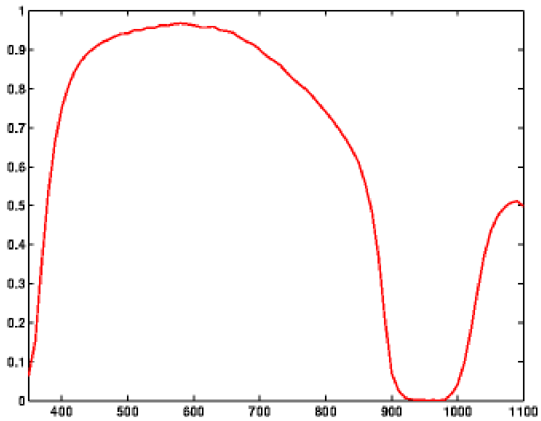

Figure 2b Lens percent transmission

As is evident, the CCD response maximum occurs at about 750nm, falling off sharply below about 550nm and lens transmission also falls off sharply below 420nm. Camera electronics are used to adjust the response of the signals through the R,G,B filters so that the signal output Pi corresponds to the correct relative strengths in the R,G,B bands, compensating for non-linearities in the CCD response and lens transmission above. From the above, the "blue" channel amplifier gain is set higher than the red or green bands due to the truncation and decrease in CCD responsivity at short wavelengths as shown, potentially introducing noise error in the blue.

These two factors are significant for accurate colour representation and image measurements, as wavelengths below 500nm greatly affect the CIE tristimulus X and Z coordinate values and to a lesser extent Y, which are needed for robust and accurate colour measurements.

For the reasons outlined above, the light source chosen for the reference measurements has a correlated colour temperature close to 6500 Kelvin, corresponding to λ(max) of 444nm, satisfying the spectral criteria for colour measurement under the above constraints.

2c Calculation of reference matrix

A number of commercially available invariant* reference colours were measured on a colorimeter [9], returning X,Y,Z tristimulus values, vector q below, under specific conditions.

A linear transformation matrix "A" was determined such that:

$$[A]\{p\}=\{q\}$$

Equation (3)

where A is the matrix ($a_{i,j}$), i={1,2,3} and j={1,2,3} and p=RGB input vector and q=XYZ vector.

* "invariant" in this context means colorimeter measurements are guaranteed to be within given tolerance under set conditions.

A corollary of work by Horn [10] is that a valid linear transform can be achieved between an input (RGB) vector p and the reference vector q (X,Y,Z), provided that the set of vectors Pi are linearly independent, and the 3x3 matrix A is non-singular. This is true for the system described with light source and detector constraints above. Careful choice of reference colours avoids non-singular matrix.

Matrix [A] was determined in a statistically "over-determined" way by using more than the required minimum of three reference colours. This technique results in reducing coefficient errors using minimisation of residual least squared errors. An added bonus of using this technique lies in immediate identification of reference colours which may have altered their colorimeter measured response, as can occur for example if a colour sample subsequently gets dirty.

Taking the inverse of matrix [A] results in a reference transform matrix herein called matrix [C], a fixed transform.

$$[C]_{fixed} = [A]^{-1}$$

Equation (4)

In operation, an input image {p} with unknown external factors of lighting, relative geometry or detector characteristics and which contains at least 3 reference colours is captured. Matrix [A] ***for this image*** is found and applied, yielding vector q. (At this point, other colour space co-ordinate transforms may be applied, for example the CIE Lab or HSI can be easily found, as all colour space co-ordinate spaces are defined in terms of the CIE XYZ.) The saved fixed transform [C] is then applied; resulting in a final referenced image {p ref} which may be compared to subsequent images.

```
{p ref} = [A][C]{p}
```
Equation (5)

3. Results

The 33 subjects were imaged using the reference system described above. These form the basis of quantifying skin colour for facial recognition purposes and make up a facial colour “truth data” set. Two colour space transforms are used below (HSV and Lab) which show a clustering of points, corresponding to the three designated groups. Group1: {Australian & Northern European}. Group2: {Sri Lankan, Indian & Egyptian}. Group3: {Chinese, Vietnamese & Taiwanese}.

It is evident that Group 1 and Group 3 are essentially inseparable as many points occupy the same region and there is significant boundary overlap of these 2 classes.

The class made up of Group 2 instances are clearly separable wrt Groups 1 and 3. Each point on the graph represents the modal value of a homogeneous area on an image greater than 800 pixels, where homogeneity here is defined by comparing the mean and mode and standard deviation (RGB) values over each area. Figure 3 below.

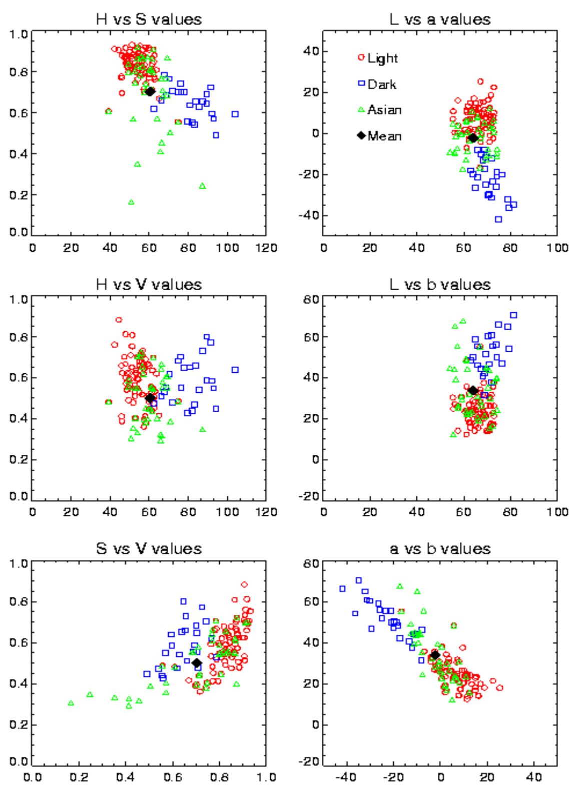

Figure 3 Showing robust cluster results of Group 1 Light (red) , Group 2 Dark (blue), Group 3 Asian (green) & mean (black).

Plots using the original R,G,B bands exhibit a scattered response, no clustering was apparent.

Some examples of pre & post corrected images and typical results of simple segmentation follow.

Original image 1a under unknown illumination, facial colour is too crimson, corrected image 1b shows correct colour rendition (light brown) per condition 1 above.

Image 1a original image

Image 1b corrected colours

Image 1 segmentation

Original image 2a under low colour temperature lighting. Face colour too red, corrected image 2b.

Image 2a original

Image 2b corrected

Image 2 segmentation

Image 3a under old fluorescent lighting too blue, corrected image 3b.

Image 3a original

Image 3b corrected

4. Limitations

Useable RGB image values range between 10% and 95% of full scale. Values less than 10% of full scale generally lead to quantisation errors. For image values close to saturation, highly non-linear behaviour due mainly to camera CCD response or camera electronics or frame-grabber characteristics may be unreliable.

Uniform illumination both vertically and across an image is desirable, as is fixed local geometry and stable characteristics of camera output.

5. Conclusions

Common schemes using face skin colour as a means of identifying or segmenting faces in complex surroundings rely on a pre-defined definition of 'skin-colour'. These are usually only robust when environmental characteristics are well controlled. The technique described corrects these unknown factors in a robust way, retaining the inherent tri-dimensionality of the colour data and improving on earlier techniques outlined in [2],[3],[4],[5],[6]. Results show good results using simple segmentation.

An interesting result of this work is the ability to subdivide all facial colours from the 8 nationalities tested into 2 major groups, where the "Light & Asian" classes form cluster 1 and the "Dark" class forms cluster 2. This paper has described an effective means of compensating for external factors affecting the colour in an image, resulting in a robust identification of skin colour.

References.

[1] P N Hall, E Claridge, and J D M Smith, Computer screening for early detection of melanoma - is there a future? British Journal of Dermatology, vol. 132, pp. 325-338, 1995. PubMed Abstract: http://www.ncbi.nlm.nih.gov/htbinpost/Entrez/query?uid=7718448&form=6&db=m&Dopt=r

[2] H Frey and H Palus, Sensor calibration for video-colorimetry, in Proceedings on Workshop on Design Methodologies for Microelectronics and Signal Processing, Gliwice,Cracow,Poland, October 1993,pp.109-113.

[3] M Herbin, F X Bon, A Venot, F Jeanlouis, M L Dubertret, L Dubertret, and G Strauch, Assessment of healing kinetics through true color image processing, IEEE Transactions on Medical Imaging, vol.12, no.1, pp.39-43, Mar 1993.

[4]M Herbin, A Venot, J Y Devaux, and C Piette, Colour quantitation through image processing in skin, IEEE Transactions on Medical Imaging, vol. 9, no. 3, pp. 262-269, 1990.

[5] S Gammal and R Popp, A color image analysis system (cd-cwa) to quantify wound healing of ulcers, Skin Research and Technology, vol. 1, no. 3, pp. 158, 1995.

[6] C Solomon, A R Munro, A M van Rij, and R Christie, The use of video image analysis for the measurement of venous ulcers, British Journal of Dermatology, vol. 133, pp. 565-570, 1995. PubMed Abstract: http://www.ncbi.nlm.nih.gov/htbin-post/Entrez/query?uid=7577585&form=6&db=m&Dopt=r

[7] Sangwine, S. J., Colour in Image Processing, Electronics & Communication Engineering Journal(5), October 2000, 211-219.

[8] Merritt, T.P. & Hall, F.F., Blackbody Radiation, Proceedings of the IRE, pp1435-1441, 1959, paper 3.1.2.

[9] J. Beresford, "Instrumental measurement of colour", J. Oil Colour Chem. Association, pp800-820, 1970, paper 53.

[10] Horn, B.K.P., Exact Reproduction of Colored Images, Journal of Computer Vision, Graphics, and Image Processing 26, pp135-167, 1984.

[11] G.Wyszecki & W.S. Stiles, "Color Science", John Wiley & Sons, New York, 1982.

[12] J.C. Russ, "The Image Processing Handbook" 3rd edition, CRC Press, Florida, 1998.