Color Calibration for a Dermatological Video Camera System

C. Grana, G. Pellacani, S. Seidenari, R. Cucchiara
University of Modena and Reggio Emilia
{grana.costantino, pellacani.giovanni, seidenari.stefania, cucchiara.rita}@unimore.it

Abstract

In this work we describe a technique to calibrate images for skin analysis in dermatology. Using a common reference we correct non-uniform illumination effects, give an estimation of the gamma correction, produce a XYZ conversion matrix. The final result is then reverted to a non standard RGB color space, built from the instrument images. In this way different instruments behave uniformly allowing colorimetric characterization, while improving the results of common algorithms. The proposed techniques should be the initial support for a distributed framework where dermatological images can be consistently compared.

1. Introduction

Color analysis has proved to be an important factor in the diagnostic process of dermatoscopic images, and therefore degradation and errors have shown to impact on the diagnostic ability of the clinician. Not much investigation has been conducted on the consistent interchange of dermatoscopic images and on the differences in color description between different instruments. Even if images can appear roughly similar in successive comparisons, the problem of non consistent color representation becomes critical in computer-based automatic color analysis. For example, clinicians are used to manually tune the light intensity so that images appear *bright enough*; in this case any comparison on the dark to light variation (for example in the segmentation of dark areas) is influenced by this setting, posing a serious question if it is possible to use those images at all in an automatic framework. Over this, the colors found on enlarged skin images cover only a small subspace of the visible colors, much smaller then that of standard RGB color spaces, and this produces a too coarse representation when working with 8 bit per channel images.

We need a common reference to compare results of different instruments and an assessed unique procedure of calibration and color representation that must be adopted in all acquisitions. We do not propose a new color space, but we define a technique that can be used, potentially for whichever acquisition system, to calibrate it with respect to the XYZ system and to produce a known transformation to a specific RGB space, especially describing that instrument. This would enable the exchange of images in distributed medical DBMS, along with their color space definition to easily and consistently compare images from different instruments.

2. Related work

Many papers, for example [1,2,3], deal with the use of digital dermoscopy for melanoma diagnosis, and use color information to assess the presence of colors typical of malignant lesions. However, the same effort wasn't put in the models and techniques for consistent image acquisition, leaving with the problem of determining how generalizable the obtained results are. Moreover, there is no agreement on the best choice for color space in skin image analysis.

The most complete approach to color calibration for the dermatological field was described by Haeghen in [4], in which a complete setup and a calibration phase of a specifically defined instrument was thoroughly described. In this paper we follow the outline of Haeghen's work, extending the model to adjust the calibration to commercial instruments and proposing a method to build a transformation that allows the clinicians to keep using their machinery, but enabling them to produce images with better characteristics of reproducibility and standardization.

3. Instrument calibration

The first necessary step of color calibration is the correction of the not uniform illumination of the instrument. The basic assumption here is that if we are imaging a uniform reference surface, we should obtain a nearly uniform reading for the whole image. This is not true because of

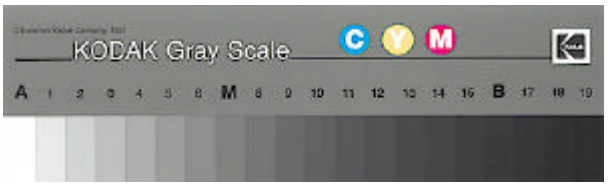

Figure 1 – The Kodak Gray Scale used for filter and gamma estimation.

the instrument characteristics, but we can measure the deviation from uniformity to invert it, thus obtaining a correction map for all the acquired images. A masking process must also occur to get rid of pixels that do not convey any data, like top or bottom lines that are always black because of the frame grabber settings, or the black ring that some instruments present at lower magnification levels.

To measure the color deviation, the filter values are computed separately for each color channel: calling $V_C(x,y)$ the value of the channel C at the pixel of coordinates (x,y), the value of the filter F_C is:

$$F_C(x,y) = \frac{V_C(x,y) - M_C}{M_C} \tag{1}$$

where M_C is the most represented value in a selected window of the image. This value is easily computed from the histogram of the channel, as the bin with maximal value:

$$h_C(i) = \#\{(x,y) : C(x,y) = i\} \tag{2}$$

$$M_C = \underset{i \in [0,255]}{\operatorname{argmax}}\, h_C(i) \tag{3}$$

The window selection should be taken in a well lit area to uniform the image to the zone that is best illuminated. This allows using M_C also as an indicator of channel saturation. In fact it should be carefully avoided to obtain clipped values during the calibration phase, since these could strongly affect the final result. If a channel has saturation effects, it is no more possible to describe it with linear transforms.

Moreover this filter estimation has to be conducted in a way that guarantees independence from the particular uniform surface that we selected and from the specific acquisition. So we used the Kodak Gray Scale (Fig. 1) as a reference for our calibration. After computing the M_C^i and F_C^i for each of the 20 steps of the gray scale, and discarding those squares whose M_C^i values clip at 255, we evaluate $\tilde{F}_C$ as:

$$\tilde{F}_C(x,y) = \underset{j \in [0,19]}{\operatorname{argmin}} \sum_{i=0}^{19} \left| F_C^i(x,y) - F_C^j(x,y) \right| \tag{4}$$

that is the median of the measured values. Each image I is then filtered for each channel using the equation:

$$\tilde{I}_C(x,y) = \frac{I_C(x,y)}{1 + \tilde{F}_C(x,y)} \tag{5}$$

obtaining the filtered image $\tilde{I}$.

3.1. Gamma estimation

After obtaining light compensated images, the next step is to estimate the non linear relation between the luminance factor, also known as CIE tristimulus value Y, and the digital values provided by the camera. Y is linearly related to incident light and is a standard of light measurement. Moreover, given fixed light source and geometry, it should be linearly related to the reflectance R of the surface, that can be estimated from the optical density OD as $R = 10^{2-OD}$.

Measuring the Kodak Gray Scale with our instrument (Fig. 2) we effectively verified on its outputs the presence of a gamma correction of the RGB values, that is a power relation between the digital measures and the known Y values [5].

To give an estimate of this relation we compared the normalized values of $\hat{Y} = Y/Y_n$ and $d_C = V_C/255$, $C = R,G,B$ with the following equation:

$$y = ax^g + b \tag{6}$$

and estimated the three parameters (a,b,g) separately for the three channels. The aim was to use the N measured values d_i, with reference to the currently considered color channel, and the corresponding $\hat{Y}_i$ to obtain the triplet:

$$(a,b,g) = \underset{a,b,g}{\operatorname{argmin}} \sum_{i=1}^{N} \hat{Y}_i - \left(a d_i^g + b\right). \tag{7}$$

The estimation was conducted by exhaustive search in a ±0.5 space around an initial guess point (1.0, 0.5, 0.5) with 0.1 steps. After finding the triplet that produced the minimum value, the search was repeated at finer steps around that point, narrowing progressively the search

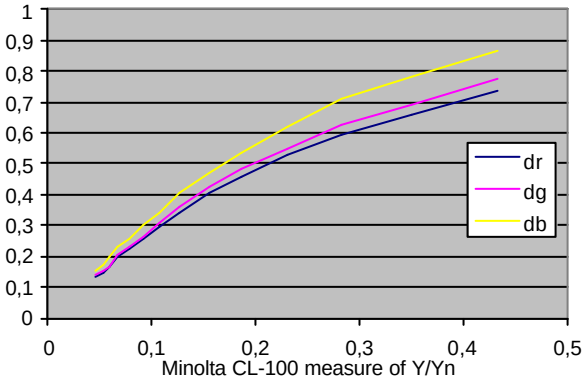

Figure 2 - Measured values for gamma correction estimation.

	a	b	?
R	1.8088	-0.6009	0.2903
G	1.8580	-0.5859	0.3019
B	2.0493	-0.6439	0.2982

Table I – Estimated values for the gamma correction of each channel of the instrument.

space, up to the desired precision. In case the solution was at any limit of the defined space, the search was repeated without narrowing the area, just moving the center to the best point found. Table I shows the values obtained for the gamma correction of the instrument under analysis.

3.3. Transformation matrix estimation

After the gamma estimation, it is easy to invert the power relation and obtain triplets of RGB values that can be linearly transformed by matrix multiplication into XYZ triplets. The question is how to do this conversion, that is which matrix to use. To answer this question, another reference object is used: the GretagMacbeth Color-Checker Color Rendition Chart, a target that is often used in broadcast television in order to evaluate the color accuracy of broadcast television cameras. It is produced using painted papers, so it is not the perfect surface for a skin reference, but its matte surfaces allow for easier handling of reflexes and the squares are large enough for quick positioning of the measuring instrument. This target also was measured with the Minolta CL-100 and the resulting values were found quite consistent with tabulated data, given along with the target.

Each square of the ColorChecker was acquired and filtered with the previously estimated filter, then the average value of a central rectangle of sides half the width and half the height was computed as the measured value for the patch. The maximum histogram value for each channel M_C was evaluated to reveal patches that couldn't be correctly imaged by the instrument. Differently from other works, we didn't try calibrating the instrument so that it could correctly describe all the colors in the ColorChecker, instead we chose to just evaluate the settings that the typical clinicians would use. So we discarded all the patches that presented $M_C = 0$ or $M_C = 255$.

In our experiments, 9 patches couldn't be used because of the excessively narrow range of the blue channel that saturated to 0 or 255 in all the excluded patches and because of the red channel that saturated to 255 in the three yellow/orange patches (Fig. 3).

Following a well known procedure for optimization [6], we took the M valid triplets of RGB values, their corresponding XYZ values and searched for the matrix $\boldsymbol{A}$ that gave the "best" transform in the form:

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} \cong \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} R_i \\ G_i \\ B_i \end{bmatrix} \tag{8}$$

for all the patches. This means that we solve the linear system given by:

$$\begin{bmatrix} & & & & & \cdots & & & \\ R_i & G_i & B_i & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & R_i & G_i & B_i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & R_i & G_i & B_i \\ & & & & & \cdots & & & \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{32} \\ a_{33} \end{bmatrix} = \begin{bmatrix} \cdots \\ X_i \\ Y_i \\ Z_i \\ \cdots \end{bmatrix} \tag{9}$$

In most cases, this can be solved exactly for three patches, but this leads to a perfect conversion for the three colors and a terrible one for the others. To solve with more patches we used the Singular Value Decomposition (SVD) that gives the best solution in the sense that the solution gives the minimum difference between the known XYZ and those estimated from the RGB values. Other objectives could be searched for, as for example to minimize the difference measured in the CIE $L^*a^*b^*$ color space, ΔE or CIE94 [6]. Haeghen has shown experimentally that no evident benefit was obtained in dermatological images by proceeding this way, even if this approach requires more difficult and thus less common minimization algorithms.

The relation between RGB and XYZ values is not always well described by a linear transform, so to cope also with slightly more complex relations, we used a non-linear operator that included the covariance terms:

$$\Theta_9 \begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} R & G & B & RG & GB & BR & R^2 & G^2 & B^2 \end{pmatrix}^{\mathrm{T}} \tag{10}$$

This operator was described by Haeghen and provided the best fit with the ColorChecker data, without overfitting. So, after applying this operator to the data, matrix $\boldsymbol{A}$ becomes 3 rows by 9 columns, but nothing changes in the way we proceed with optimization.

At this point we have obtained an estimation of the XYZ color coordinates for every measured pixel. The aim is to transform from this values to another known color space that allows the visualization and a simpler storage. In [4] the sRGB color space was chosen because it is

Figure 3 – Simulated GretagMacbeth ColorChecker. The left one shows the theoretical values in the sRGB color space, while the right one is the measured one with the saturating patches.

based on the phosphors used in many modern CRT-based display devices, including computer monitors. A problem arises in color space conversions, and it is due to the quantization that has to be applied to the values. In fact this causes to divide a certain "light" range into a fixed number of steps (without being interested in the effect of gamma correction on the steps' characteristics). The problem with this fixed division is that if we are interested in a limited area of the sRGB color space, we lose much detail that is set aside to describe color space regions that our images will never use. For all these reasons, we propose to describe the instrument color space extracted as the best inverse transform that reproduced correctly a set of original images of the instrument chosen as training set. This average characterization is built to get a better use of the color representation in the spectrum area occupied by this kind of images. Its known relation with XYZ, will allow an easy conversion to any other chosen color space, and the practice of viewing the images on an sRGB calibrated computer monitor (a common setting available in most modern monitors), will allow for a common evaluation of images, even if obtained from different sources. Our proposed conversion uses a 3x3 matrix to supply a linear conversion and applies a gamma conversion. The fact that this color space was extracted to describe a specific instrument requires more investigation to see if it is suitable also for other instruments.

4. Experimental results

In this work we used the FotoFinder video microscope by TeachScreen Software GmbH, which consists in a video camera with an illuminating head with microscopic lenses. The system allows acquisition of lossless compressed images with a resolution of 767x576 at 24 bit. It is heavily subjected to heating effects and needs about one hour of warm up time, to get sufficiently stable readings. To obtain an acceptably stable reading during time, we had to insert pauses between series of acquisitions, so that the instrument had the time to assess. A series of acquisitions on the same uniform surface, taken with 1 minute distance and with a pause of 10 minutes after every 6 captures lead to a variation from the first to the last acquisition of less than 3 levels per channel on average, that seemed more than enough for our purposes.

After the calibration of two different FotoFinder equipments, we acquired the same training set of lesions with both instruments and converted the images in the color space estimated from the first instrument. Apart from a visual inspection, by which the resulting images showed much more conformance, we tried to compare a simple histogram measure to get a verification of the results. A median cut algorithm was run on the lesion and a 4 bin

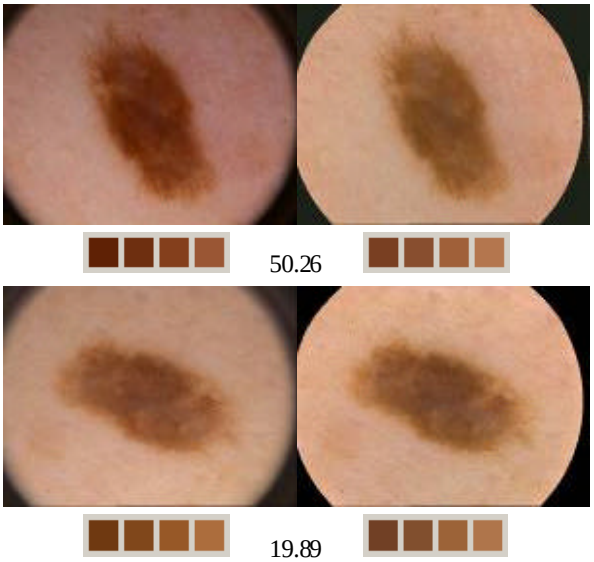

Figure 4 – Example of two images of different instruments (left side) and their conversion to a common colorspace. The distance between the two histograms is shown.

description in the final RGB color space was obtained. The squared distance between corresponding color bins was used. An example of image comparisons is provided in Fig. 4. Using our proposal an average reduction of the difference between images acquired by two instruments was more than the 60%.

5. Conclusions

We described a practical technique to calibrate images for the dermatological context. No particular instruments are needed, save the use of a common reference, and the final result can still be reverted to a defined RGB color space able to well fit the specific instrument used. Even not colorimetric aware techniques (as most in this area literature) can benefit from the described workflow.

6. References

[1] J. Chen, R.J. Stanley, R.H. Moss and W.V. Stoecker, "Colour analysis of skin lesion regions for melanoma discrimination in clinical images", Skin Res. Tech., vol. 9, 2003, pp. 94-104.

[2] R. Cucchiara, C. Grana, S. Seidenari, G. Pellacani, "Exploiting Color and Topological Features for Region Segmentation with Recursive Fuzzy c-means", Machine Graphics and Vision, vol. 11, no. 2/3, 2002, pp. 169-182.

[3] S.E. Umbaugh, R.H. Moss, W.V. Stoecker, G.A. Hance, "Automatic Color Segmentation Algorithms: With Application to Skin Tumor Feature Identification", IEEE Engineering in Medicine and Biology, vol. 12, no. 3, Sept 1993, pp. 75-82.

[4] Y.V. Haeghen, J.M.A.D. Naeyaert, I. Lemahieu, and W. Philips, “An Imaging System with Calibrated Color Image Acquisition for Use in Dermatology”, IEEE Transactions on Medical Imaging, vol. 19, no. 7, Jul 2000, pp. 722-730.
[5] R.S. Berns, “The Science of Digitizing Paintings for Color-Accurate Image Archives: A Review”, Journal of Imaging Science and Technology, vol. 45, no. 4, Jul/Aug 2001, pp. 305-325.
[6] R.S. Berns, *Billmeyer and Saltzman’s Principles of ColorTechnology*, 3rd ed. Wiley-Interscience, New York, 2000.