Colorimetric Object Classification

Wolfram Hans Dietrich Paulus
Universität Koblenz-Landau
Computational Visualistics
{hans,paulus}@uni-koblenz.de

Abstract

In order to improve object recognition results, usually several image preprocessings are performed. If color images are used, a color normalization is normally applied. Algorithms for color normalization will be compared to a colorimetric approach found in the literature. Recovering colorimetric values instead of a simple RGB camera output leads to more reliable color images. To this kind of processed object image a basic object recognition approach using different histogram distances is applied. It will be shown that there is an effect on the results of object recognition rates if we use color calibrated images instead of color normalization methods.

Keywords: Object recognition, color normalization, color calibration

1 INTRODUCTION

The use of color histograms as features is widely used to solve the object recognition task and is described in detail e.g. in [11]. In In our contribution, the color of the object is not the main aspect of interest but the statistical distribution of the image itself. We use the RGB value and count it in a histogram bin, rather than recording the *color* defined by a colorimetric tristimulus value. We use three histograms per image that are given by each of the R, G, and B color channel in the RGB value case as well as in the colorimetric case.

The RGB values formed by the camera depends heavily on the image formation process - especially the illumination involved. Mainly for this reason color normalization algorithms are applied to estimate the influence of pose and color of the illumination and eliminate – or at least minimize – their influence to the image appearance. In several situations such color normalizations leads to an improvement of the recognition rates that are presented in section 5. The question arises whether the use of calibrated color values in the histograms lead to another raise of these rates.

The following section gives an overview to some familiar color normalization algorithms. The calibration approach used in the experiments is presented in section 3. Results are discussed in sections 4 and 5.

2 COLOR NORMALIZATION

Using a normalized color one might expect the object features to be better distinguishable. Hence numerous color normalization approaches have been proposed (e. g. [1, 3, 9, 7, 2]).

One very simple method normalizes the color of one pixel (r, g, b) by dividing it by the lightness $(r + g + b)$. The results are called chromaticities.

A very well known assumption is the world to be gray. It was formalized by Buchsbaum [1] and postulates in average the color of a natural scenes sums up to a gray.

A combination of the aforementioned was proposed by Finlayson [3] who iteratively applies the two ideas above in his so-called 'comprehensive color image normalization' (CCN) algorithm.

Another idea that uses the gray world assumption is the color normalization by rotation. All color values are considered to be coordinates in a three dimensional color space. All pixel together form a color cloud with a preferred direction. Rotating their principal vector onto the main diagonal, which is the gray axis of the color space, leads to a normalized representation. Several color spaces are used by different authors. One that uses the RGB color space is proposed by Paulus [9].

Since color constancy is a nature of human perception, some methods for color normalization try to copy this feature. One well known approach is called 'Retinex' and was introduced by Land [7]. A modification used for the experiments is done by Fankle [4].

The Retinex is also capable to normalize the effect of different light sources, while most other algorithms assume only one. One approach that tries to deal with local changes in illumination is called 'local space average color' (LSAC) and was introduced by Ebner [2] who assumes that the changes are moderate within the scene. The impact of several color normalization algorithms is shown in Figure 1. Upper left shows the original image from the capturing device, bottom right illustrates the colorimetric calibrated version.

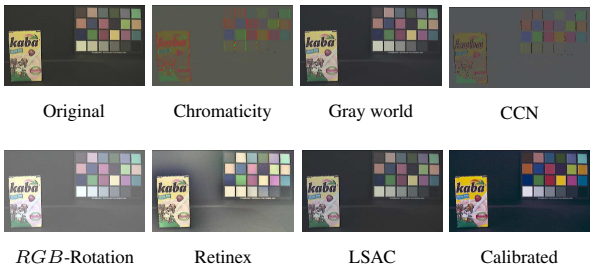

Figure 1: Influence of several color normalization methods used in the experiments to the unprocessed image (top left). Bottom right the result of the calibration method proposed by Lee [8].

3 CALIBRATION

Using a calibration for cameras leads to more reliable results concerning the color of an object. Besides a photometric calibration that determines the brightness transfer function, a radiometric calibration is mandatory for colorimetric information. We use a a method that provides a colorimetric calibration as proposed by Lee [8].

3.1 Image formation

A simplified model of image formation considers the light source, the reflectance of the object, and the cameras sensor sensitivity. The brightness transfer function is assumed to be equal to one, making the relation between radiance and the sensors response to be linear. Camera outputs that do not hold this assumption are linearized in a preprocessing step prior to the calibration.

The simplified image formation model is built as

$$f^{(k)} = \int_{\lambda} E(\lambda) \cdot \rho(\lambda) \cdot R_k(\lambda)\, \mathrm{d}\lambda \qquad (1)$$

Herein $f^{(k)}$ is used as the sensors response which is the signal at the output of the cameras channel k. Normally, a RGB-triplet value ($K = 3, k \in \{r, g, b\}$) is used. The spectral composition of the light source, denoted by $E(\lambda)$, is multiplied with the spectral reflectance $\rho(\lambda)$ of the object. Hence the cameras incident light is $E\rho$ with the related spectrum. The multiplication with the spectral sensitivity $R_k(\lambda)$ leads to the sensor response which is the integrated value over the spectral range of the sensor. The spectral calculation of the color values in the experiments is limited to the range of 380 nm to 730 nm with respect to the employed measurement device 'Eye-One Photo' of Gretag Macbeth. The width of the sampling interval is $\Delta\lambda = 10\,\mathrm{nm}$ which leads to $L = 36$ samples. The discretized version of (1) is

$$f^{(k)} = \sum_{n=1}^{L} E_\lambda \cdot \rho_\lambda \cdot R_{k,\lambda} \cdot \Delta\lambda \qquad (2)$$

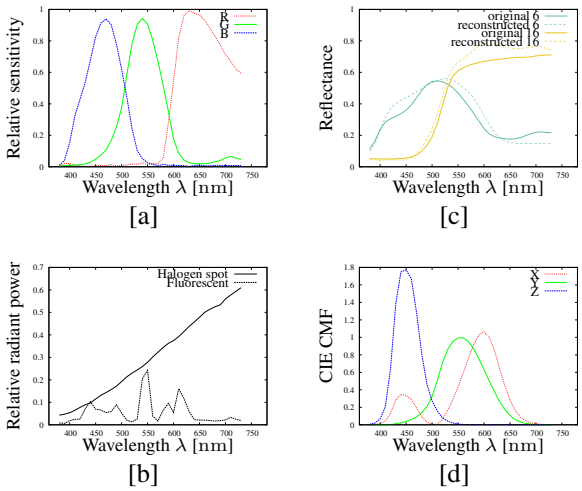

Figure 2: Spectra used for the calibration: Camera sensitivity [a], the light sources used during the image acquisition [b], exemplary two reflectance spectra of the patches 6 and 16 of the ColorChecker© with their reconstructed spectra received by the applied calibration [c], as well as the CIE color matching functions used to get the color values XYZ [d].

or rewritten in a matrix notation $\boldsymbol{f} = \boldsymbol{PER}$ where $\boldsymbol{f}$ denotes the $1 \times K$ output vector, $\boldsymbol{P}$ the $1 \times L$ reflectance vector, $\boldsymbol{E}$ the $L \times L$ illumination spectrum (a matrix with E_{λ_l} on the diagonal), and $\boldsymbol{R}$ the $L \times K$ sensor sensitivity matrix. The $\Delta\lambda$ as a constant implicitly is contained in the sensor response.

3.2 Colorimetric calibration

A precondition for calibration as described in Lee [8] is a controlled environment to set the parameters. After calibration, a colorimetric tristimulus can be calculated from RGB values.

The first step is to determine the cameras sensitivity curves. Lee requires for his method an initial estimate $\hat{\boldsymbol{R}}\,(L \times K)$ but gives no hint about the necessary quality. For this reason the manufacturers specifications (Figure 2[a]) are used. This first approximation is improved by a 3×3 correction matrix $\tilde{\boldsymbol{R}}$ that contains an adaption to the prevailing acquisition situation. For this reason an image of the ColorChecker© is captured under a known illumination (2[b]). The coefficients of the spectral reflectance of the color patches, (e.g. Figure 2 [c]), are known from the measurements with the photospectrometer, too. The $Q = 24$ RGB-triple that complies with the camera response for each patch is merged into a $Q \times K$ matrix $\boldsymbol{F}$. The unknown correction matrix $\tilde{\boldsymbol{R}}$ is determined by a Moore-Penrose pseudoinverse of $\boldsymbol{F} = \boldsymbol{PE}\hat{\boldsymbol{R}}\tilde{\boldsymbol{R}}$. The new, adapted camera sensitivity $\boldsymbol{R}$ is given by multiplication of $\hat{\boldsymbol{R}}$ and $\tilde{\boldsymbol{R}}$. One has to keep in mind that this matrix $\boldsymbol{R}$ has a bias induced by the RGB values of the ColorChecker© and the illumination $\boldsymbol{E}$.

???

The next step is to perform a PCA on the spectra of the calibration target. Lee states, that the reconstruction of a spectrum could be done sufficiently correct by three basis vectors that are combined in a $b \times L$ matrix $\boldsymbol{P}_{\mathrm{B}}$. The associated weights are called $\boldsymbol{G}$ $(1 \times b)$. Together with the mean spectrum of all patches $\boldsymbol{P}_{\mathrm{M}}$, $(1 \times L)$ the original spectra can be reconstructed by

$$\boldsymbol{P} = \boldsymbol{P}_{\mathrm{M}} + \boldsymbol{G}\boldsymbol{P}_{\mathrm{B}}. \tag{3}$$

With this, (2) can be rewritten as

$$\boldsymbol{G} = (\boldsymbol{f} - \boldsymbol{P}_{\mathrm{M}}\boldsymbol{E}\boldsymbol{R})(\boldsymbol{P}_{\mathrm{B}}\boldsymbol{E}\boldsymbol{R})^{-1}. \tag{4}$$

This is the central equation for reconstructing a spectrum by calculating the weight $\boldsymbol{G}$ as a function of the RGB values $\boldsymbol{f}$. Putting the weight into (3) leads to a reconstructed spectra.

Since the values are only correct for the patches, in the following several correction steps are performed to get colorimetrically correct values. That is why a $Q \times K$ matrix $\boldsymbol{C}$ is introduced. For each channel (columns) and patch (rows) the element is determined by the theoretical color values given by $\boldsymbol{P}\boldsymbol{E}\boldsymbol{R}$ divided by the measured RGB values in $\boldsymbol{F}$. An additional correction is carried out by multiplying each row of $\boldsymbol{C}$ with the corresponding ratio of the sum of the reconstructed spectrum to the sum of the measured spectrum. Using $\boldsymbol{f} = \boldsymbol{C} \circ \boldsymbol{F}$ leads now to a colorimetrically correct RGB-value where '$\circ$' denotes the element by element multiplication.

Still, RGB-values that are not part of the 24 ColorChecker© patches has to be modified to result in a good spectral estimation. The data has to be interpolated, assuming smooth spectral changes leads to smooth variations in the RGB values.

Hence the spectrum of an unknown triple can be reconstructed by their neighbouring values from know patches. This is done by the Euclidean distance from the $\boldsymbol{f}_{\mathrm{r}}$ (that has to be reconstructed) to the every element (patch) of $\boldsymbol{F}$. This distance is multiplied by a weight factor h to reduce the influence of patches with a larger distance. Lee propose h to be $\ln(10^{-9})$. For the distance to the j^{th} patch $W(j) = e^{\frac{h \cdot D[j]}{\mathrm{range}}}$. These elements are part of the $1 \times Q$ 'inverse distance weighting vector' $\boldsymbol{W}$. Range means the difference between the minimum and maximum value the camera is capable of, usually 255 using 8 bit. The experiments are performed by using the minimum and maximum value within the image while assuming good-natured content.

When normalizing $\boldsymbol{W}$ by its sum and multiplying by $\boldsymbol{C}$, the result is close to one for those values produced by the patches and zero to those not examined but not exactly one, or zero respectively. This is due to the interpolation performed. For this reason another $L \times L$ correction matrix $\boldsymbol{W}'$ is applied. Its j^{th} row contains the vector $\boldsymbol{W}$ described above for the respective patch. To maintain the original correction factor for the ColorChecker© patches a final correction matrix $\boldsymbol{C}'$ $(L \times K)$ is introduced: $\boldsymbol{C}' = \left(\boldsymbol{W}'\right)^{-1}\boldsymbol{C}$. Using a final correction vector $\boldsymbol{c} = \boldsymbol{W}\boldsymbol{C}'$ to modify the RGB-value $\boldsymbol{f}_{\mathrm{r}}$ that has to be reconstructed to get the values $\boldsymbol{f}$ inserted in (4). The resulting weights are used in (3) to reconstruct the spectrum of which two samples are shown in Figure 2 [c].

This spectral estimations are assessed with the color matching functions (Figure 2 [d]) to create the tristimulus values XYZ which we transform to the RGB-color space in our experiments.

4 EXPERIMENTS

The examinations are based on images taken from KOPID[1] that contains 17 objects recorded under three different illuminations and at 12 varying viewpoints. Additionally the camera was moved to five diverse levels of height. Unlike other image databases such as COIL-100[2] or ALOI[3] the KOPID-images contain the ColorChecker© what makes it suitable for a colorimetric calibration based on known color samples.

Other than the histogram comparisons in [11] the histograms in the experiment at hand are created for each RGB-channel separately, i.e. we use three one-dimensional histograms. This simplifies the calculation of the Earth Mover's Distance distance (EMD) [10] which we use for histogram comparsion. The information loss compared to the three-dimensional histograms is justifiable by the results which we show below.

The images were subject to one of the color normalizations mentioned above. The quantization of the histograms is done from 8 to 16 bins per channel. As distance measures the sum of squared differences (SSD), histogram intersection (HI), a χ^2 distance, and the EMD are used. The recognition rates are determined as follows: The histogram of the object illuminated by the fluorescent spectrum is compared to all other histograms of the objects illuminated by the halogen spot. If the smallest difference in the histograms is found in an image containing the same object, the recognition task is counted as successful. The recognition rate is calculated by the number of successfully classified objects divided by the total amount of tests performed.

5 RESULTS

In Figure 3 the advance of the EMD (match distance respectively) compared to the other histogram distance measurements is obvious. The rates obtained by the SSD are the least promising. For this reason, the results

[1] http://www.uni-koblenz.de/kopid
[2] http://www1.cs.columbia.edu/CAVE//software/softlib/coil-100.php
[3] http://staff.science.uva.nl/~aloi/

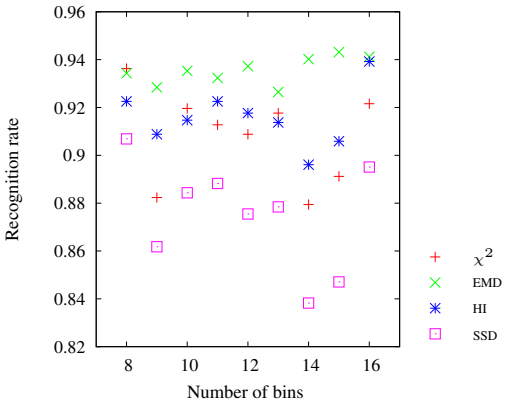

Figure 3: Object recognition rates while using different histogram distance measures. The set used for this case contains the calibrated images only.

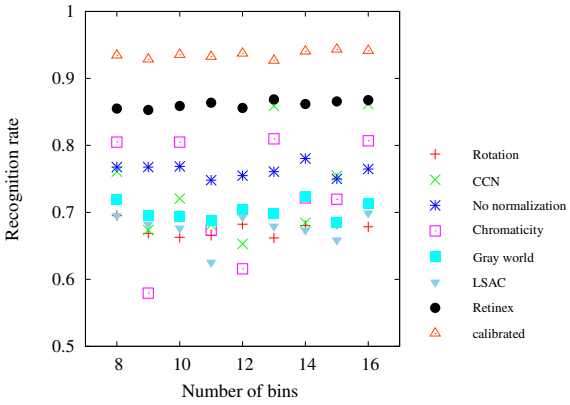

Figure 4: Results using different color normalization algorithms. The blue stars determine the recognition rates where no color normalization is applied. The histogram distance used to obtain this rates is the EMD.

presented in Figure 4 are achieved by the EMD comparison.

The first benchmarks to be reached are the rates received by no normalization (blue stars in Figure 4). These are the recognition rates without performing any color normalization.

It is obvious that some algorithms do not hold the expectation to improve the recognition rate. This is the rotation, gray world, and LSAC. The CCN and chromaticity have marks on both sides of the blue borderline. Using the arithmetic mean they still fail to meliorate the recognition rate of no normalization.

The color normalization performed by the retinex-algorithm displaces the aforementioned by $86{,}11\%$ correct classifications in average. The Matlab® implementation provided by [5] performs a normalization in 8 seconds on a 3GHz computer. That is roughly five times as fast as the colorimetric calibration used.

Nevertheless the calibration applied leads to an additional raise of the recognition rates of around 7% and performs at $93{,}54\%$ in average on the top of all normalization methods. Compared to the no normalization this is a 17% boost.

6 CONCLUSIONS

To consolidate the results in object recognition rates several extensions could be added. First, the transfer to three-dimensional histograms should be considered. Second, the impact of other color spaces than RGB has to be analyzed. Third, the partly heavy variations in some of the normalization methods could be analyzed. One idea here is to increase the number of objects in the database.

The poor performance of some of the algorithms might be a topic for further experiments with other or larger databases or different bin numbers. A larger sample size could also be contemplated for the calibration results - the knowledge of (spectral) image formation given.

Anyhow, finally the results show that a colorimetric calibration outperforms the recognition rates received by commonly used color normalization algorithms. The price of a laborious calibration is well invested into higher recognition rates.

REFERENCES

[1] G. Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310:1–26, 1980.

[2] Marc Ebner. Color constancy based on local space average color. *Machine Vision and Applications*, 2008.

[3] Graham D. Finlayson, Bernt Schiele, and James L. Crowley. Comprehensive colour image normalization. *Lecture Notes in Computer Science*, 1406:475–490, 1998. 1406.

[4] Jonathan J. Frankle and John J. McCann. Method and apparatus for lightness imaging, 1983. US Patent no. 4,384,336.

[5] Brian Funt, Florian Ciurea, and John McCann. Retinex in Matlab. In *Proceedings of IS&T/SID Eighth Color Imaging Conference*, pages 112–121, 2000.

[6] Wolfram Hans, Benjamin Knopp, and Dietrich Paulus. Farbmetrische Objekterkennung. In Gerd Stanke and Michael Pochanke, editors, *15. Workshop Farbbildverarbeitung*, pages 43–51, Berlin, Germany, 9 2009. GfAI.

[7] E.H. Land and J.J. McCann. Lightness and retinex theory. *Journal of the Optical Society of America*, 61(1):1–11, 1971.

[8] Raymond L. Lee. Colorimetric calibration of a video digitizing system: algorithm and applications. *Color Research and Applications*, 13(3):180–186, 1988.

[9] Dietrich Paulus, L. Csink, and Heinrich Niemann. Color cluster rotation. In *Proceedings of the International Conference on Image Processing (ICIP)*, Chicago, 10 1998. IEEE Computer Society, IEEE Computer Society Press.

[10] Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas. The Earth Mover's Distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2):99–121, 2000.

[11] Michael J. Swain and Dana H. Ballard. Color indexing. *International Journal of Computer Vision*, 7(1):11–32, 1991.