Robust Histogram Construction from Color Invariants

Theo Gevers
Faculty of Science, University of Amsterdam
1098 SJ, Amsterdam, The Netherlands
gevers@science.uva.nl

Abstract

A simple and effective object recognition scheme is to represent and match images on the basis of color histograms. To obtain robustness against varying imaging circumstances (e.g. a change in illumination, object pose, and viewpoint), color histograms are constructed from color invariants. However, in general, color invariants are negatively affected by sensor noise due to the instabilities of these color invariant transforms at many RGB values. To suppress the effect of noise blow-up for unstable color invariant values, in this paper, color invariant histograms are computed using variable kernel density estimators. To apply variable kernel density estimation in a principled way, models are proposed for the propagation of sensor noise through color invariants. As a result, the associated uncertainty is known for each color invariant value. The associated uncertainty is used to derive the parameterization of the variable kernel density estimator during histogram construction. It is empirically verified that the proposed method compares favorably to traditional color histograms for object recognition.

1 Introduction

Color provides powerful information for object recognition. To provide object recognition robust against the accidental imaging conditions (e.g. illumination, shading, highlights, and viewpoint), color histograms are computed from color invariants [2], [3], [12]. For example, simple and effective illumination-independent color ratio's have been proposed by Funt and Finlayson [2] and Nayar and Bolle [7]. Further, for the dichromatic reflection model, Gevers and Smeulders [3] proved that normalized color rgb ($c_1c_2c_3$) is to a large extent invariant to a change in camera viewpoint, object pose, and for the direction and intensity of the incident light. In addition, the hue color H ($l_1l_2l_3$) is insensitive to highlights under white light or white-balanced camera. However, these color invariant transforms bring with them various serious drawbacks since these transformations are singular at some points and unstable at many others. For example, color ratio's, rgb and $c_1c_2c_3$ color space become unstable near the black point while hue H and $l_1l_2l_3$ are very unstable along the achromatic axis [3]. As a consequence, a small perturbation of RGB values will cause a large jump in the transformed values introducing severe errors during histogram construction. Traditionally, the effects of noise blow-up at unstable colors are suppressed by ad hoc thresholding. For example, all pixels in a color image are discarded with a local saturation and intensity smaller then 5 percent of the total range [3]. Inevitably, more elaborated computational methods are required to construct robust histograms from color invariants.

Therefore, in this paper, a more principled method is proposed to suppress the effect of noise during histogram construction from color invariants. To this end, variable kernel density estimators are employed to construct color invariant histograms. To apply variable kernel density estimation in a proper way, models are presented for the propagation of sensor noise through color invariants. As a result, the associated uncertainty is known for each color invariant value. The associated uncertainty is used to derive the parameterization of the variable kernel density estimator used during histogram construction.

This paper is organized as follows. In section 2, related work is reviewed. A theoretical model is presented, in section 3, to derive the uncertainty associated at each color invariant value. In section 4, the optimal kernel parameterization is proposed and incorporated in the variable kernel density estimator. Then, in section 5, in the context of color-based object recognition, the variable kernel density estimator is compared to traditional histograms.

2 Previous Work

Kender [5] shows that hue and normalized color are singular at some RGB values and unstable at many others. For instance, the essential singularity of normalized coordinates is at black ($R = G = B = 0$), whereas for the hue the singularity is at the achromatic axis ($R = G = B$). As a consequence, both color spaces become unstable near the singularity where a small perturbation of RGB value might cause a large jump in the transformed values. Traditionally, these effects are either ignored or suppressed by ad hoc thresholding of the transformed values. For example, Ohta [8] rejects normalized color values if the sum of RGB is less than 30, and rejects hue values if the saturation times the intensity is less than 9. Healey [4] rejects all sensor measurements that fall within the sphere of radius 4σ centered at the origin in RGB space. In order to indicate how the color space transform influences the mean, variance and covariance of the colors under the influence of noise, Burns and Berns [1] analyze the error propagation from a measured color signal to the CIE $L^*a^*b^*$ color space. Shafarenko, Petrou and Kittler [9] use an adaptive filter for noise reduction in the CIE $L^*u^*v^*$ space prior to 3-D color histogram construction. The filter width depends on the covariance matrix of the noise distribution in the CIE $L^*u^*v^*$ space.

As opposed to previous work, the aim of this paper is to use kernel density estimators (see e.g. [10, 14]) to construct color invariant histograms. To apply variable kernel density estimation in a principled way, the error propagation of sensor noise is analyzed through the color invariants. As a result, the associated uncertainty is known for each color invariant value. Then, kernel sizes are adapted with respect to the amount of noise blow-up for unstable color invariant values. As a consequence, unstable color invariant values will contribute less to the final histogram than stable color invariant values. Although our method of variable kernel density estimation is suited for different color invariants, in the sequel, we focus on normalized color rgb ($c_1c_2c_3$) and hue as these color spaces are widely in use in computer vision tasks.

3 Error Propagation Through Color Invariants

The result of a measurement of a quantity u is stated as:

$$\hat{u} = u_{\text{est}} \pm \sigma_u \tag{1}$$

where u_{est} is the best estimate for the quantity u (the average value) and σ_u is the uncertainty or error in the measurement of u (the standard deviation). Suppose that $u, \cdots, w$ are measured with corresponding uncertainties $\sigma_u, \cdots, \sigma_w$, and the measured values are used to compute the function $q(u, \cdots, w)$. If the uncertainties in $u, \cdots, w$ are independent, random and relatively small, then the predicted uncertainty in $\hat{q}$ [13] is:

$$\sigma_q = \sqrt{\left(\frac{\partial q}{\partial u}\sigma_u\right)^2 + \cdots + \left(\frac{\partial q}{\partial w}\sigma_w\right)^2} \tag{2}$$

where $\partial q/\partial u$ and $\partial q/\partial w$ are the partial derivatives of q with respect to u and w.

In any case, the uncertainty in q is never larger than the city block distance

$$\sigma_q \leq \left|\frac{\partial q}{\partial u}\right|\sigma_u + \cdots \left|\frac{\partial q}{\partial w}\right|\sigma_w \tag{3}$$

Based on the measured RGB-values, the normalized color rg is computed by:

$$\begin{aligned} r &= R/(R+G+B) \\ g &= G/(R+G+B) \end{aligned} \tag{4}$$

Substitution of (4) in (2) gives the uncertainty for the normalized coordinates

$$\begin{aligned} \sigma_r &= \sqrt{\frac{R^2(\sigma_B^2+\sigma_G^2)+(B+G)^2\sigma_R^2}{(B+G+R)^4}} \\ \sigma_g &= \sqrt{\frac{G^2(\sigma_B^2+\sigma_R^2)+(B+R)^2\sigma_G^2}{(B+G+R)^4}} \end{aligned} \tag{5}$$

where σ_R, σ_G and σ_B denote the uncertainties in RGB-space. In Appendix A, a noise model is given to derive σ_R, σ_G and σ_B. Further, c_1c_2 color space is given by:

$$\begin{aligned} c_1 &= \arctan R/G \\ c_2 &= \arctan B/G \end{aligned} \tag{6}$$

Again substitution of (6) in (2) gives the uncertainty for the c_1c_2 coordinates

$$\begin{aligned} \sigma_{c_1} &= \sec^2 \frac{RG\sigma_R^2 - R^2\sigma_G^2}{G^3} \\ \sigma_{c_2} &= \sec^2 \frac{BG\sigma_B^2 - B^2\sigma_G^2}{G^3} \end{aligned} \tag{7}$$

The hue θ is computed as

$$\theta = \arctan\left(\frac{\sqrt{3}(G-B)}{(R-G)+(R-B)}\right) \tag{8}$$

Substitution of (8) in (2) gives the uncertainty for the hue as

$$\begin{aligned} \sigma_\theta &= \frac{1}{2}\sqrt{3}\left(\frac{-2BR\sigma_G^2 + R^2\sigma_B^2 + \sigma_G^2 + G^2(\sigma_B^2+\sigma_R^2)}{B^2+G^2-GR+R^2-B(G+R)^2}\right) \\ &+ \frac{1}{2}\sqrt{3}\left(\frac{B^2(\sigma_G^2+\sigma_R^2) - 2G(R\sigma_B^2 + B\sigma_R^2)}{B^2+G^2-GR+R^2-B(G+R)^2}\right) \end{aligned} \tag{9}$$

4 Histogram Construction by Variable Kernel Density Estimation

A density function f gives a description of the distribution of the measured data. Perhaps the best known density estimator is the histogram. The (one-dimensional) histogram is defined as

$$\hat{f}(x) = \frac{1}{nh} \text{(number of } X_i \text{ in the same bin as } x\text{)} \quad (10)$$

where n is the number of pixels X_i in the image, h is the bin width and x the range of the data. Two choices have to be made when constructing a histogram. First, the binwidth parameter needs to be chosen. Secondly, the position of the bin edges needs to be established. Both choices effect the resulting estimation. Alternatively, the kernel density estimator is insensitive to the placement of the bin edges. The formula is

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) \quad (11)$$

Here, kernel K is a function satisfying $\int K(x)dx = 1$. In the *variable* kernel density estimator, the single h is replaced by n values $\alpha(X_i), i = 1, \cdots, n$. This estimator is of the form

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\alpha(X_i)} K\left(\frac{x - X_i}{\alpha(X_i)}\right) \quad (12)$$

The kernel centered on X_i has associated with it its own scale parameter $\alpha(X_i)$, thus allowing different degrees of smoothing. To use variable kernel density estimators for color images, we let the scale parameter be a function of the RGB-values and the color space transform. We are now left with determining the scale and shape of the kernel.

In most experiments, when the number of measurements increases, the histogram takes a definite shape. According to the noise model given in Appendix A, the limiting distribution for measuring photons is Poisson. When the average number of counts is large, the Poisson distribution is approximated well by the Gauss distribution [13]. We therefore define the shape of the kernel by the normal distribution

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp^{\frac{-x^2}{2}} \quad (13)$$

The variable kernel method estimates the univariate, directional hue density as

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^{n} \sigma^{-1} K\left(\frac{(x - \theta) \bmod (2\pi)}{\sigma}\right) \quad (14)$$

where σ is derived according (9). The variable kernel method for the bivariate normalized rgb kernel is given by:

$$\hat{f}(x, y) = \frac{1}{n} \sum_{i=1}^{n} \sigma_r^{-1} K\left(\frac{x - r}{\sigma_r}\right) * \sigma_g^{-1} K\left(\frac{x - g}{\sigma_g}\right) \quad (15)$$

where $*$ denotes convolution and σ_r, σ_g are derived according (5). Further, the bivariate normalized c_1c_2 kernel is given by:

$$\hat{f}(x, y) = \frac{1}{n} \sum_{i=1}^{n} \sigma_{c_1}^{-1} K\left(\frac{x - c_1}{\sigma_{c_1}}\right) * \sigma_{c_2}^{-1} K\left(\frac{x - c_2}{\sigma_{c_2}}\right) \quad (16)$$

where $*$ denotes convolution and $\sigma_{c_1}, \sigma_{c_2}$ are derived according (7).

In conclusion, to reduce the effect of sensor noise during density estimation, we use variable kernels where the normal distribution defines the shape of the kernel. Further, kernel sizes are adapted with respect to the amount of noise blow-up for unstable color invariants. As a consequence, the influence of unstable color invariant values will be less than stable color invariant values.

5 Experiments

In this section, the performance of the proposed variable kernel density estimator will be evaluated. Firstly, the sensitivity of the density estimator is tested with respect to noise. Then, our method is experimentally verified and compared with traditional histogram schemes in the context of color based object recognition. For the experiments, the hue range is defined from $0°$ to $360°$ over $1°$ interval. The normalized color range is defined from 0 to 255 units over 1 unit intervals. The images are obtained using a Sony XC-003P color camera and Matrox Corona framegrabber.

5.1 Propagation of Uncertainties

Th aim of this experiment is to empirically verify the validity of the proposed model of noise propagation through color invariant formulae. To this end, for the hue color space θ, the uncertainty σ_ϕ is computed for each color pixel according to (9). Further, the actual (measured) uncertainty is derived as the standard deviation of hue values averaged over an homogeneously colored region as recorded by the color camera. In this way, the experiment is conducted on nine images taken from homogeneous colored sheets of paper material. Sheet number 1 has a bright red color, number 2 is red colored, 3 is yellow, 4 is lightgreen, 5 is green, 6 is cyan, 7 is darkblue, 8 is blue and 9 is purple. From this experiment, we obtain that the absolute difference between the predicted (computed by (9)) and actual hue values is $0.07^\circ \pm 0.03^\circ$ which is well below 1 percent of the hue range. Further, the experiment is repeated for normalized colors where the uncertainty is computed according to (5). The absolute difference between the predicted and measured normalized red values is 0.4 ± 0.2 and for the green values is 0.2 ± 0.1 which is again well below 1 percent of the normalized color range. The experiment shows that the computed (predicted) uncertainties compare favorably to the measured (actual) uncertainties.

5.2 Sensitivity of Density Estimators for Unstable Color Values

The goal of this experiment is to evaluate the performance of the density estimation methods for dark and bright colors. Therefore, two images are taken of a variable density filter placed on top of a yellow and blue colored paper sheet. The filter attenuates an incident beam of radiation without altering its spectral distribution. The intensity varies across the filter's length. As a result, one part of the image is bright, the other part is dark. The number of local maxima present in the density estimations is taken as a measure for the sensitivity to sensor noise. The results are given in Table 1. The uncertainty in the hue for dark colors is larger than for bright colors as a result of the instability of hue. The experiment shows that the number of local maxima present in the kernel density estimation compares favorable to the estimation by the histogram method.

5.3 Color Based Object Recognition

In this section, we consider object recognition on the basis of color invariant histograms. Therefore, in section 5.3.1, the dataset is presented. Then, in section 5.3.2, we compare traditional histogram-based object recognition with variable kernel density estimation.

5.3.1 Dataset

For comparison reasons, the same dataset used by [3], [11], has been taken to conduct the experiments. These images are recorded by the SONY XC-003P CCD color camera and the Matrox Magic Color frame grabber. Two light sources of average day-light color are used to illuminate the objects in the scene. The database consists of $N_1 = 500$ target images taken from colored objects, tools, toys, food cans, art artifacts etc. Objects were recorded in isolation (one per image). The size of the images are 256x256 with 8 bits per color. The images show a considerable amount of shadows, shading, and highlights. A second, independent set (the query set) of $N_2 = 70$ query or test recordings was made of randomly chosen objects already in the database. These objects were recorded again one per image with a new, arbitrary position and orientation with respect to the camera, some recorded upside down, some rotated, some at different distances.

Then for each image, traditional histograms (c.f. eq. 10) and variable density estimators are constructed on the basis of rg and θ. For the histograms, we have determined the appropriate bin size for our application empirically by varying the number of bins on the axes over $q \in \{2, 4, 8, 16, 32, 64, 128, 256\}$. The results show (not presented here) that the number of bins was of little influence on the retrieval accuracy when the number of bins ranges from $q = 32$ and up. Therefore, the color histogram bin size for each axis used during histogram formation is $q = 32$.

5.3.2 Robustness against Noise

We study the performance of the variable kernel estimator with respect to noise for the 70 test images and 500 target images. For comparison reasons in the literature, matching is based on histogram intersection [12].

For a measure of match quality, let rank r^{Q_i} denote the position of the correct match for test image Q_i, $i = 1, ..., N_2$, in the ordered list of N_1 match values. The rank r^{Q_i} ranges from $r = 1$ from a perfect match to $r = N_1$ for the worst possible match.

Then, for one experiment, the average ranking percentile is defined by:

$$\bar{r} = \left(\frac{1}{N_2}\sum_{i=1}^{N_2}\frac{N_1 - r^{Q_i}}{N_1 - 1}\right)100\% \qquad (17)$$

The effect of noise is produced by adding independent zero-mean additive Gaussian noise with $\sigma \in \{2, 4, 8, 16, 32, 64\}$ to the query images. In Figure 1, two images are shown generating together 10 images by adding noise with $\sigma \in \{8, 16, 32, 64, 128\}$.

We concentrate on the quality of the recognition rate on the basis of hue θ with respect to different noise levels. Further, to compare traditional histogram matching with the proposed kernel density estimator, we have constructed four different histograms. Firstly, no thresholding has been performed. This histogram construction scheme does not cope with unstable color invariant values. Hence, all color invariant values are equally weighted in the histogram, as used by [12]. The color histogram without thresholding is denoted by $\mathcal{H}_{\theta_1}$ based on the hue θ color model. Secondly, we have thresholded θ values if the intensity is below 5 percent of the total range as proposed by [3, 8]. For this histogram construction scheme, we denote $\mathcal{H}_{\theta_2}$ to be based on θ. Thirdly, only hue values are considered during histogram construction if the intensity and saturation are within a range of 4σ centered at the origin of the RGB space, used by [4], yielding $\mathcal{H}_{\theta_3}$. Finally, histogram are based on the proposed variable kernel density estimator given by $\mathcal{H}_{\theta_4}$.

Color	Hue	Histogram #Max.	Kernel #Max.
Bright yellow	225 ± 2	2	1
Dark yellow	210 ± 40	46	27
Bright blue	68 ± 1	2	1
Dark blue	56 ± 9	16	3

Table 1. *Comparison of performance of density estimation methods for dark (unstable) and light (stable) colors in hue color space. The average number of local maxima present in the estimates and is taken as a measure for sensitivity to unstable hue values.*

Figure 1. *2 images generating together 10 images by adding noise with* $\sigma \in \{8, 16, 32, 64, 128\}$.

The influence of noise differentiated by the various histogram construction schemes, shown in Figure 2 based on the hue color model, shows that kernel density estimator outperforms the ad hoc thresholding schemes. In fact, the kernel density estimator gives good results up to considerable amounts of noise ($\sigma = 64$). Further, the thresholding schemes give always higher recognition accuracy than no thresholding at all.

6 Conclusion

In this paper, variable kernel density estimation is used to construct robust color invariant histograms. The variable kernel density estimation is derived from a theoretical framework for noise propagation through color invariants. In this way, the associated uncertainty is computed for each color invariant value which is used to steer the kernel sizes. From the theoretical and experimental result we conclude that kernel density estimator overcome the problem of ad-hoc thresholding at unstable color invariants. Further, our method is far less sensitive to Gaussian noise than traditional histogram construction schemes.

Appendix A: Model for the prediction of sensor noise

Modern CCD cameras are sensitive enough to be able to count individual photons. Photon noise arises from the fundamentally statistical nature of photon production because a different number of photons will be counted for two different recordings during the same time interval t. The probability distribution for counting ρ photons during T seconds is known to be Poisson. For a Poisson distribution, the average number of photons measured by an image sensing element at position (x, y) is given as

$$f(x, y) = \rho t \pm \sqrt{\rho t} \tag{18}$$

where ρ is the rate of photons measured per second. Our aim is to analyze how the uncertainty $\sigma_f = \sqrt{\rho t}$ propagates to the uncertainty in transformed color coordinates.

To determine the uncertainty associated with a measured gray value, the number of electrons, or sensitivity S, necessary to change from one brightness level to the next is given by [6] as

$$S = 1/\gamma \tag{19}$$

The sensitivity can be measured as follows: After compensating for the dark-current offset, the camera output is given by

$$g(x, y) = \gamma \cdot f(x, y) \tag{20}$$

Substitution of the measured number of photons and the associated uncertainty of (18) in (20) gives

$$g(x, y) = \gamma \cdot (\rho t \pm \sqrt{\rho t}) \tag{21}$$

Using the mean $\overline{g} = \gamma \cdot \rho t$ and standard deviation $\sigma_g = \gamma \cdot \sqrt{\rho t}$ of a region with homogeneous pixel values, the sensitivity S for a particular CCD camera is derived as

$$S = \frac{\overline{g}}{(\sigma_g)^2} \tag{22}$$

From (21) follows that the uncertainty in the number of photons measured at an arbitrary pixel $g(x, y)$ is given by

$$\sigma_p(x, y) = \sqrt{\frac{g(x, y)}{S}} \tag{23}$$

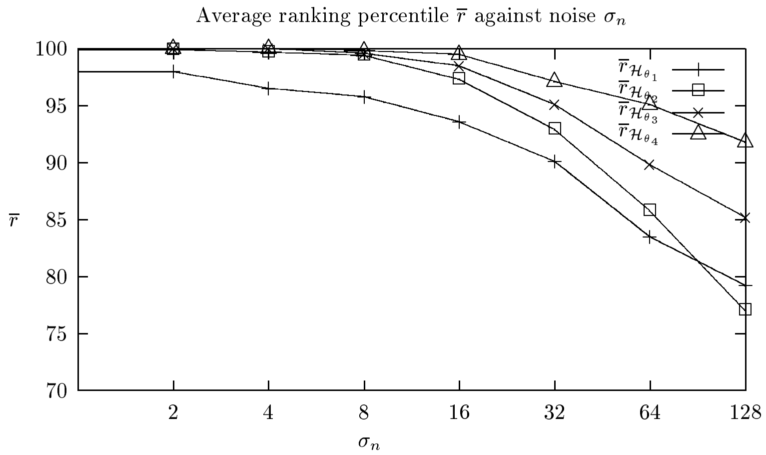

Figure 2. *The discriminative power of the matching process differentiated for the various histogram construction schemes based on θ with respect to noise. The average percentile $\overline{r}$ for histogram $\mathcal{H}_{\theta_1}$, $\mathcal{H}_{\theta_2}$, $\mathcal{H}_{\theta_3}$ and $\mathcal{H}_{\theta_4}$ is given by $\overline{r}_{\mathcal{H}_{\theta_1}}$, $\overline{r}_{\mathcal{H}_{\theta_2}}$, $\overline{r}_{\mathcal{H}_{\theta_3}}$, $\overline{r}_{\mathcal{H}_{\theta_4}}$ respectively.*

Under the assumption that the dark-current noise σ_d and photon noise σ_p are independent, the total amount of noise σ_n for a certain pixel value is

$$\sigma_n(x, y) = \sqrt{\sigma_d^2 + \sigma_p^2} \tag{24}$$

References

[1] P. D. Burns and R. S. Berns, "Error propagation analysis in color measurement and imaging," *Color Research Applications* **22**, August 1997.

[2] B. V. Funt and G. D. Finlayson, Color Constant Color Indexing, *IEEE Trans. PAMI* **17(5)**, 522-529, (1995).

[3] T. Gevers and A. W. M. Smeulders, "Color based object recognition," *Pattern Recognition* **32**, pp. 453–464, March 1999.

[4] G. Healey, "Segmenting images using normalized color," *IEEE Transactions on Systems, Man and Cybernetics* **22**(1), pp. 64–73, 1992.

[5] J. R. Kender, "Saturation, hue and normalized color: Calculation, digitation effects, and use," tech. rep., Carnegie-Mellon University, November 1976.

[6] J. C. Mullikin, L. J. van Vliet, H. Netten, F. R. Boddeke, G. van der Feltz, and I. T. Young, "Methods for ccd camera characterization," in *Image Acquisition and Scientific Imaging Systems*, H. C. Titus and A. Waks, eds., vol. 2173, pp. 73–84, SPIE, 1994.

[7] S. K. Nayar, and R. M. Bolle, Reflectance Based Object Recognition, International Journal of Computer Vision, Vol. 17, No. 3, pp. 219-240, 1996

[8] Y.-I. Ohta, T. Kanade, and T. Sakai, "Color information for region segmentation," *Computer Graphics and Image Processing* **13**, pp. 222–241, 1980.

[9] L. Shafarenko, M. Petrou, and J. Kittler. Histogram-based segmentation in a perceptually uniform color space. *IEEE Transactions on Image Processing*, 7(9):1354–1358, September 1998.

[10] B. W. Silverman, *Density Estimation*, Chapman & Hall, 1986.

[11] N. Sebe and M.S. Lew and D.P. Huijsmans, "Toward improved ranking metrics", *IEEE Trans. on PAMI*, **22(10)**, pp. 1132-1143, 2000.

[12] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision* **7**(1), pp. 11–32, 1991.

[13] J. R. Taylor, *An Introduction to Error Analysis*, University Science Books, 1982.

[14] M. P. Wand and M. C. Jones, *Kernel Smoothing*, Chapman & Hall, 1995.