Probabilistic invariant image representation and associated distance measure

Jorge Scandaliaris
DEEC, FACET
Universidad Nacional de Tucumán
Tucumán, Argentina
jscandaliaris@herrera.unt.edu.ar

Alberto Sanfeliu
Automatic Control Department
Universitat Politècnica de Catalunya (UPC)
Barcelone, Spain
sanfeliu@iri.upc.edu

Abstract

Varying illumination is a limiting factor for many computer vision applications, especially in outdoor settings. Invariant image representations aim to reduce this effect and provide the following processing steps, image segmentation, edge detection, object recognition, etc., with a more stable view, closer to the surface reflectances presents in the scene than to the illumination. In this work we present an invariant image representation that integrates several key observations in a probabilistic way and an associated probabilistic distance measure. Used together, they can be used as a measure of similarity between the surfaces represented by a given pair of pixels, even under illumination color changes.

1. Introduction and related work

The log-chromaticity [2] space has some interesting properties regarding illumination color changes. Under the assumption of Planckian illumination, Lambertian surfaces and narrow band camera sensors, it can be shown that illumination color changes translate into shifts along a camera-dependent direction for all surfaces. This particular property was first pointed out by Finlayson et al. [2], where they used it to create an invariant image representation

Several works have studied the variation of natural light and showed that typical range, expressed in the temperature of an equivalent blackbody radiator or correlated color temperature, lies approximattely between 5000 to 15000 K. In [7] we took into consideration the limited variation of natural light to improve upon the results obtained with Finlayson's intrinsic image for tracking in an outdoor setting. The feature space used for tracking was the log-chromaticity space, and the initial model was enlarged in the direction of illumination color change by smoothing with an anisotropic Gaussian filter. The results suggests that the method retained the invariant properties while at the same time increasing the discriminative power when compared to Finlayson's intrinsic image. In this work, in addition to the shifting effect and the limited variation of natural illumination described above, we also include the effect of sensor noise, which translates into uncertainties being dependent on intensity due to the non-linear nature of the log-chromaticity space, and integrate them in a probabilistic way. The result is a probabilistic model were knowledge about sensor uncertainty and plausible variations of illumination color are integrated into the log-chromaticity space and combined with a probabilistic distance measure.

The paper is organized as follows: In Section **??**, we review the log-chromaticity space. In Section **??**, we model the illumination color variation. In Section **??**, we propagate sensor noise into the log-chromaticity space and then in Section **??**, we describe the proposed model. The probabilistic distance measure is introduced in Section **??**.

2. Log-chromaticity space

Forming the 3-vector chromaticities by dividing each band by the geometric mean, $\sqrt[3]{RGB}$, and then calculating their logarithm, we arrive at

$$\rho_k = \ln\left(\frac{R_k}{(\prod_{i=1}^{3} R_i)^{1/3}}\right),\ k = 1, 2, 3 \qquad (1)$$

where R_k denote the sensor responses. It can be shown [3] that under the assumptions of Plackian illuminants, Lambertian surfaces and narrow band camera

sensitivities, $\underline{\rho}$ has the form

$$\underline{\rho} = \ln \underline{s} + \frac{1}{T}\underline{e} \qquad (2)$$

where $\underline{s}$ depends on the surface and the camera, $\underline{e}$ is independent of the surface, but which again depends on the camera, and T is the illuminant color temperature. All 3-vector $\underline{\rho}$ lie on a plane orthogonal to $\underline{u} = 1/\sqrt{3}(1,1,1)^\mathsf{T}$. The redundant dimension is removed by transforming 3-vectors $\underline{\rho}$ into a coordinate system *in* the plane using a 2×3 matrix $\boldsymbol{U}$, (see [3] for details)

$$\underline{\chi} \equiv \boldsymbol{U}\underline{\rho}, \ \underline{\chi} \ is \ 2 \times 1 \qquad (3)$$

Equation (2) is the parametric form of a line. Given that $\underline{e}$ is independent of the surface, the lines defined by all surfaces have the same slope. Illumination color changes, affecting T, shift points along the lines defined by each surface. This behavior is retained in (3). Athough this conclusion is based on restrictive assumptions, it has been shown [3] to be a good approximation in real situations.

3. Illumination color variation

We model natural illumination by a black body radiator, and more precisely by Wien's approximation to Planck's law for the spectral radiance of a black body. Wien's approximation has the form

$$E(\lambda, M) = \frac{2hc^2}{\lambda^5} \exp\left(-\frac{hcM}{\lambda k}\right) \qquad (4)$$

where, by convenience, we express T in reciprocal mega-Kelvin, $M \equiv \frac{10^6}{T}$. k is Boltzmann's constant, c is the speed of light, h is Planck's constant, and M has units MK^{-1}. The spectral power distribution, that is the color, of the illumination is determined by the parameter M. We model M as having a normal distribution,

$$M \sim \mathcal{N}(\mu_M, \sigma_M^2) \qquad (5)$$

This assumption is a simplification of the observed distribution in experimental data (see, for example, [4], but reduces the complexity of our model.

4. Propagation of sensor noise into the log-chromaticity space

Models for the noise in CCD cameras [9] include a term related to thermal noise, independent of the amount of light arriving at the sensor, and another, shot noise, dependent on it. Thermal noise is more important at low light levels, while shot noise dominates at higher intensities. Then, noise statistics will vary across the image, depending on the local intensity level. For simplicity, we assume that sensor noise is normally distributed, $n(x,y) \sim \mathcal{N}(0, \sigma_{x,y})$.

It can be shown [8] that if q is a function of several quantities $x, \ldots, z$ measured with uncertainties $\sigma_x, \ldots, \sigma_z$, then

$$\sigma_q = \sqrt{\left(\frac{\partial q}{\partial x}\sigma_x\right)^2 + \cdots + \left(\frac{\partial q}{\partial z}\sigma_z\right)^2} \qquad (6)$$

Assume now that we have determined the standard deviation of the sensor responses, σ_R, σ_G and σ_B, for each pixel of the image. The two components of $\underline{\chi}$, can be expressed in terms of R, G and B, after some algebraic manipulations, as

$$\chi_1 = \frac{\sqrt{2}}{2}(\ln R - \ln B) \qquad (7)$$

$$\chi_2 = \frac{\sqrt{6}}{6}(-\ln R + 2\ln G - \ln B) \qquad (8)$$

Substituting (7) and (8) into (6) we arrive at the expressions for the standard deviation of the two components of the log-chromaticity space, σ_{χ_1} and σ_{χ_2} we get

$$\sigma_{\chi_1} = \sqrt{\frac{\sigma_B{}^2}{2B^2} + \frac{\sigma_R{}^2}{2R^2}} \qquad (9)$$

$$\sigma_{\chi_2} = \sqrt{\frac{\sigma_B{}^2}{6B^2} + \frac{2\sigma_G{}^2}{3G^2} + \frac{\sigma_R{}^2}{6R^2}} \qquad (10)$$

Lower intensity levels in the RGB image result in higher standard deviations in the log-chromaticity space.

5. Probabilistic model and distance measure

The components of Equation (2) are

$$\chi_1^i = k1\ln(s_1) + Me_1 \qquad (11)$$

$$\chi_2^i = k2\ln(s_2) + Me_2 \qquad (12)$$

where $k1$ and $k2$ are constants, and the s_k depend on the surface and the camera, the e_k depend on the camera, and M is the illumination color temperature in mired, and we use the superscript in χ_k^i denotes that these are ideal values. As stated in the previous section, we model M as a Gaussian random variable. For a given camera the e_k terms are constant, and for a given surface and camera the s_k terms are also constant. This means that χ_1^i and χ_2^i are linear transformations of M. The family of normal distributions is closed under linear transformations [6]. If X is a Gaussian random variable with mean μ_X and variance σ_X and if Y is a linear

transformation of X, such that $Y = aX + b$, for some real numbers a and b, then Y is normally distributed, and its mean and variance are $a\mu_X + b$ and $a^2\sigma_X^2$, respectively. Coming back to χ_1^i and χ_2^i, their distributions are

$$\chi_1^i \sim \mathcal{N}\left(e_1\mu_M + \ln\left(s_1\right), e_1^2\sigma_M^2\right) \quad (13)$$

$$\chi_2^i \sim \mathcal{N}\left(e_2\mu_M + \ln\left(s_2\right), e_2^2\sigma_M^2\right) \quad (14)$$

For simplicity, we refer to the means and variances in Equations (13) and (14) as μ_k^i and σ_k^i, with $k \in \{1, 2\}$ denoting the kth component of χ^i. We now incorporate the effect of a non-ideal sensor. As stated in Section 4, sensor noise is modelled as a being normally distributed and it is propagated to the log-chromaticity space. Non-ideal sensor responses, in the log-chromaticity space, will have the form

$$\chi_1 = \chi_1^i + n_1 \quad (15)$$

$$\chi_2 = \chi_2^i + n_2 \quad (16)$$

where the $n_k, k \in \{1, 2\}$ represent sensor noise after being propagated to each component of the log-chromaticity space, $n_1 \sim \mathcal{N}(0, \sigma_{\chi_1}^2)$ and $n_2 \sim \mathcal{N}(0, \sigma_{\chi_2}^2)$. Because the χ_k^i and n_k are normally distributed, then the χ_k will also be normally distributed. Moreover, the means and variances are easily calculated [6] from the respective parameters of χ_k^i and n_k

$$\chi_1 \sim \mathcal{N}\left(\mu_1^i, \sigma_1^{i\,2} + \sigma_{\chi_1}^2\right) \quad (17)$$

$$\chi_2 \sim \mathcal{N}\left(\mu_2^i, \sigma_2^{i\,2} + \sigma_{\chi_2}^2\right) \quad (18)$$

Note that each pixel defines a random variable, and because differences in intensity and the surface imaged affect both the mean and variance of its associated random variable, we have now a set of Gaussian random variables with possibly differing means and variances.

We have thus far expanded the definition of the log-chromaticity space to include two major effects that introduce variability using a probabilistic approach. To actually make use of this new definition, we need to be able to compare pixel values including this extra information. The Kullback-Leibler (KL) divergence is a widely used measure for comparing distributions. It can be interpreted as measuring the dissimilarity between two distributions. For distributions P and Q of a continuous random variable, the KL divergence is most commonly presented [1] as

$$D_{\text{KL}'}(P\|Q) = \int_{-\infty}^{\infty} p(x) \ln \frac{p(x)}{q(x)} dx \quad (19)$$

where p and q denote the probability density functions of P and Q, respectively. It can be shown that

$D_{\text{KL}'}(P\|Q) \geq 0$, with the equality holding if and only if $P = Q$. The KL divergence as defined above is not symmetric, that is, $D_{\text{KL}'}(P\|Q) \not\equiv D_{\text{KL}'}(Q\|P)$, but one can use

$$D_{\text{KL}}(P\|Q) = \frac{1}{2}\left(D_{\text{KL}'}(P\|Q) + D_{\text{KL}'}(Q\|P)\right) \quad (20)$$

instead, as initially proposed by Kullback and Leibler [5] themselves. This form is the one adopted from now on.

The random variables representing the pixels in the log-chromaticity space are by definition Gaussian. In this particular case, the KL divergence can be calculated in a closed form from the parameters of the distributions. In the univariate case, where $P \sim \mathcal{N}(\mu_p, \sigma_p^2)$ and $Q \sim \mathcal{N}(\mu_q, \sigma_q^2)$, we have that

$$D_{\text{KL}'}(P\|Q) = \frac{1}{2}\Bigg(\ln \frac{\sigma_q^2}{\sigma_p^2} + \frac{\mu_p^2 + \mu_q^2 + \sigma_p^2 - 2\mu_p\mu_q}{\sigma_q^2} - 1 \Bigg) \quad (21)$$

and for the normal multivariate case

$$D_{\text{KL}'}(P\|Q) = \frac{1}{2}\Bigg(\ln \frac{|\mathbf{\Sigma}_q|}{|\mathbf{\Sigma}_p|} + \text{Tr}(\mathbf{\Sigma}_q^{-1}\mathbf{\Sigma}_p) + (\boldsymbol{\mu}_p - \boldsymbol{\mu}_q)^\mathsf{T}\mathbf{\Sigma}_q^{-1}(\boldsymbol{\mu}_p - \boldsymbol{\mu}_q) - d \Bigg) \quad (22)$$

where $P \sim \mathcal{N}(\boldsymbol{\mu}_p, \mathbf{\Sigma}_p)$ and $Q \sim \mathcal{N}(\boldsymbol{\mu}_q, \mathbf{\Sigma}_q)$, and d is the dimension. Substituting the parameters of the distributions into (21) or (22) we then use (20) to get the symmetric version of the KL divergence.

6. Experiments

Each pixel in the proposed model is a Gaussian random variable and we selected a measure for comparing them in a probabilistic way. We need now to determine the parameters, means and variances, in Equations (17) and (18). We will assume sensor noise and illumination color temperature variations have already been characterized and thus we know $\sigma_R(x, y)$, $\sigma_G(x, y)$, and $\sigma_G(x, y)$ for the camera in use. Given an input image, the first step consists in transforming from RGB to the log-chromaticity space using Equations (7) and (8). For the means in Equations (17) and (18) we resort to the measured values, that is, we set $\mu_k^i(x, y) = \chi_k^m(x, y), k \in \{1, 2\}$, where the χ_k^m are obtained from the input image using Equations (7) and (8). We proceed in this manner because the mean of
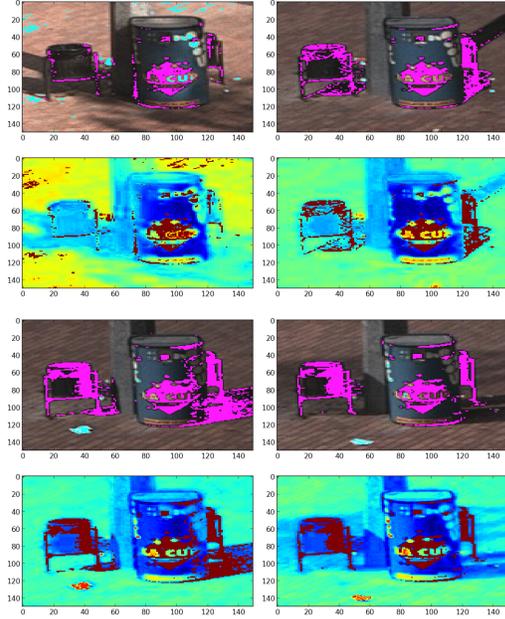
Figure 1. Example images with their associated *distance* **image. The reference pixel for all images is at** $(x,y) = (100, 60)$

the ideal responses' distribution includes a term dependent on the surface, which is unknown. The variances associated with each random variable are the sum of the variances of the illumination color change, σ_M, and the variances due to sensor noise after being propagated into the log-chromaticity space, σ_{χ_1} and σ_{χ_2}, that we determine from the σ_k, $k \in \{R, G, B\}$ and each pixel's R, G, and B values using Equations (9) and (10). In the case of M, we do not follow the model directly because for the distribution of M to be usefull, we would need to determine also the parameters in Equations (13) and (14) related to the camera, that is e_1 and e_2. Determining these parameters involves a calibration of the camera, but in return there is no need to characterize $e_k \sigma_M^2$ for each different camera. In our case we determined empirically the variance $e_k \sigma_M^2$ by imaging a set of surfaces over a wide range of illumination conditions. To determine the validity of this approach, we acquired images with a fixed camera over an extended period of time to ensure that we have a wide range of illuminations. We then calculate the distance from one pixel to the rest of the image for each image, obtaining a *distance* image. Figure **??** shows some of the images and their associated *distance* image. Note that the particular scene chosen is a worst case scenario for the method, as the chromaticities of the blue recycle bin are colinear

with those of the brick floor. Using Finalyson's intrinsic image in this scene would have shown one of its shortcomings, which is its reduce discriminative power.

7. Conclusions

We proposed a method integrating several features based on our experience using the log-chromaticity space. The key aspects exploited by our model are: the invariant properties of the log-chromaticity space, the efect of sensor noise, the limited variation of natural light, and the integration of these into a probabilistic model together with a distance measure. Even if further validation is needed, and especially with real world applications, these preliminary results are encouraging

References

[1] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1st edition, 2006.

[2] G. Finlayson and S. Hordley. Color constancy at a pixel. *J. Opt. Soc. Am. A*, 18(2):253–264, 2001.

[3] G. D. Finlayson, M. S. Drew, and C. Lu. Intrinsic Images by Entropy Minimization. In T. Pajdla and J. Matas, editors, *Proc. 8th European Conf. Comput. Vision*, volume 3024 of *Lect. Notes Comput. Sci.*, pages 582–595, Prague, May 2004. Springer-Verlag.

[4] J. Hernández-Andrés, J. Romero, J. L. Nieves, and R. L. Lee, Jr. Color and spectral analysis of daylight in southern Europe. *J. Opt. Soc. Am. A*, 18(6):1325–1335, 2001.

[5] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Statist.*, 22(1):79–86, 1951.

[6] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw Hill, New York, 3rd. edition, 1991.

[7] J. Scandaliaris and A. Sanfeliu. Discriminant and invariant color model for tracking under abrupt illumination changes. In *Proc. 20th IAPR Int. Conf. Pattern Recog.*, pages 1840–1843, Aug. 2010.

[8] J. R. Taylor. *An introduction to error analysis*. University Science Books, Sausalito, CA, 2nd edition, 1997.

[9] Y. Tsin, V. Ramesh, and T. Kanade. Statistical calibration of ccd imaging process. In *Proc. IEEE Int. Conf. Comput. Vision*, volume 1, pages 480–487, Vancouver, Canada, Jul. 2001.