Rank-Based Illumination Estimation

Lilong Shi[a] and Brian Funt[b]; [a]Advanced Image Research Lab; Pasadena, CA/USA; [b]School of Computing Science, Simon Fraser University; Vancouver, BC/CANADA

Abstract

A new two-stage illumination estimation method based on the concept of rank is presented. The method first estimates the illuminant locally in subwindows using a ranking of digital counts in each color channel and then combines local subwindow estimates again based on a ranking of the local estimates. The proposed method unifies the MaxRGB and Grayworld methods. Despite its simplicity, the performance of the method is found to be competitive with other state-of-the art methods for estimating the chromaticity of the overall scene illumination.

Introduction

The first step in automatic white balancing an image is generally to estimate the 'color' or chromaticity of the light illuminating the scene. In the camera's RGB coordinate system this is equivalent to determining the triple of (R,G,B) digital counts (up to an overall scale factor) that the camera would register from an ideal-white reflecting surface placed in the scene. Considered as a vector, it is only the direction of the (R,G,B) that matters, not its magnitude, since the magnitude is affected by the camera's exposure settings and other factors. As a result, most illumination estimation methods return the rg-chromaticity of the illuminant (r,g) where r=R/(R+G+B) and g=G/(R+G+B). Given a measured 'true' illuminant chromaticity (r_t,g_t) and an estimated chromaticity (r_e,g_e), the error in the estimate is frequently reported as the angle in degrees between the vectors (r_t,g_t,1-r_t-g_t) and (r_e,g_e,1-r_e-g_e).

There are many illumination estimation methods reported in the literature [1-12]. In this paper, we describe a new rank-based method that unifies the Grayworld and MaxRGB approaches in a manner that is analogous to, but different from, the Shades of Grey method and compare its performance to several of the better known methods for estimating the overall scene illumination. Although simple, the rank-based method performs as well or better than many of the more complicated methods.

Rank-based Method

The proposed method consists of two steps: (*i*) within local subwindows the R, G, and B digital counts within each channel are ranked separately and then the triple of digital counts of rank k_L from each channel is returned as the illuminant color for the subwindow; and (*ii*) global ranking of the subwindow estimates from (*i*) returning the R, G, and B having rank k_G. The first step estimates the illuminant locally within many different *M*-by-*M* subwindows. The RGB estimate for the central pixel of the subwindow is its k_L^{th} smallest R, k_L^{th} smallest G, and k_L^{th} smallest B. Expressed in terms of the largest instead of the smallest, the k_L^{th} smallest digital count is the (M^2- k_L) largest digital count.

When k_L is normalized by the size of the subwindow, the ranking operation is equivalent to a non-linear percentile filter (e.g., Matlab's *prctile* or *quantile* functions). For example, if the k_L^{th} rank corresponds to the 50th percentile then this non-linear filter becomes the standard median filter. For $k_L = M^2$ (100th percentile) the filter is equivalent to choosing the maximum. Hence, as k_L is varied from the 50th to the 100th percentile, the method of estimating the illumination varies from using the median, which is similar to the mean used in Grayworld estimation, to the maximum, which is equivalent to MaxRGB without pre-processing [1, 2].

The second step combines the results of the subwindow estimates into a single estimate of the illumination for the image as a whole. For a given global rank, k_G, the RGB estimate for the entire image is its k_G^{th} smallest R, k_G^{th} smallest G, and k_G^{th} smallest B found across all the subwindow estimates. The rank k_G can be normalized by the total number of subwindows to yield a percentile rank. In the remainder of the paper, k_L and k_G will always refer to normalized, percentile ranks.

Our Rank-based algorithm can also be represented as two mathematical formulas. The first formula computes the local estimates in the image, as

$$\mathbf{I}_L = Percentile_{M\text{x}M}(\mathbf{I}, k_L) \quad (1)$$

where the input image **I** is of one of the R, G or B color channels, and $\mathbf{I}_L$ contains the local illumination estimates, where each pixel is replaced by the k_L local percentile. The function PercentileMxM(x, k) returns the kth smallest value of a vector x containing MxM elements. The second formula computes the global estimates based on the local estimates as

$$e = Percentile_N(\mathbf{I}_L, k_G) \quad (2)$$

Here *e* is the estimated global illumination of one of the R, G and B color channels. The function $Percentile_N(\mathbf{x}, k)$ computes the k^{th} smallest value of a 1x*N* vector **x** (given *N* is the number of pixels in the image). So the scene illumination, (e_R,e_G,e_B), is the final output of our algorithm.

Rank Computation

Generally, to find the k^{th} smallest (or largest) element in a subwindow of M^2 elements by brute-force based on first sorting them would have complexity O(M^2log*M*). However, a much more efficient algorithm developed by Huang [13] based on using an incremental histogram has computational complexity O(*M*). Recently, Weiss [14] further reduced median filtering to sub-linear complexity O(log*M*) and a modified version performs general rank-order filtering. Therefore, for *N M-by-M* subwindows the computational complexity of the local rank step is O(*N*log*M*).

For the second step, the global ranking is based on N samples so its complexity is O(log*N*). The combined complexity for the two steps is O(*N*log*M*) +O(log*N*), however, since O(*N*log*M*) >>

O(log$N$) for large N the overall complexity of the method is O(Nlog$M$).

Comparison to other illumination estimation methods

The rank-based approach is also closely related to illumination estimation by Grayworld [3] and MaxRGB [2] and its variants [1]. MaxRGB with simple pre-processing has been shown [1] to perform at roughly the same level as the Edge-Based method [4] and significantly better than the computationally intensive Bayesian method [7] when tested on the Grayball and the Colorchecker image datasets. The proposed rank-based method unifies Grayworld and MaxRGB within a common framework in that varying the choice of 3 parameters—subwindow size M, local rank k_L, and global rank k_G—defines a family of illumination estimation methods. For instance, setting k_L = 50%, k_G = 100%, and M = 3, the rank-based method is equivalent to MaxRGB with pre-processing by a 3x3 median filter. Setting k_G = 50% and M = 1 approximates the Grayworld method with mean replaced by median.

The way that the ranked-based framework defines a family of methods is similar to the Shades-of-Gray [6] framework in that it also defines a family of methods from Grayworld through to MaxRGB; however, the two frameworks generate different families. The rank-based method also has a lot in common with the bright pixel approach of Vaezi Joze et al. [10] and Tominaga et al. [11] in that for high percentile ranks estimates are primarily based on 'bright' pixels. Table 1 summarizes the existing methods that can be implemented or approximated within the rank-based framework.

Table 1. Summary of the connection between the rank-based approach and other color constancy methods.

Parameters of Rank			CC Methods
M	k_L	k_G	
1	--	50%	Grayworld (approximated) [3]
∞	50%	--	
1	--	100%	MaxRGB without pre-processing [2]
∞	100%	--	
5	>>1	100%	MaxRGB with uniform averaging (Barnard et al.) (approximated) [15]
5	50%	100%	MaxRGB w. median filter (MaxM) [1]
>5	50%	100%	MaxRGB clipped removal + median filter (MaxCM) (approximated) [1]

Tests and Discussion

In this section, the performance of the rank-based approach is evaluated and compared with other methods on four benchmark image datasets. The first dataset is the Barnard et al. [16] collection of 321 indoor images taken under 11 different illuminants. The second is the Ciurea et al. [17] SFU dataset of 11,346 images derived from digital video sequences. The third dataset is Shi's[1] re-processed version of the Gehler et al. [9] Colorchecker dataset containing 568 images. The fourth dataset contains HDR images of 105 scenes captured using a Nikon D700 digital still camera [1]. The images in the Barnard, Gehler and HDR datasets are linear (gamma =1). The Ciurea set is non-linear and of unknown gamma.

Tests of the rank-based method were conducted using various parameter settings: (1) three subwindow sizes are defined as a fraction of the image size—approximately 0.01%, 0.1%, 1%—plus a window of size 1x1 as the limiting case; (2) local rank k_L is varied from 50% ~ 100% (i.e., from median to maximum); and (3) the global rank k_G is varied from 50%~100%, in steps of 5% down to 2% for both k_L and k_G .

The choice of parameter settings will clearly affect the method's performance as measured in terms of the angular difference between the estimated versus true chromaticity of the illumination. Fig. 1 shows the effect when tested on the 11,346 SFU Grayball image set in which the colored map in k_L-k_G coordinates indicates the magnitude of the median angular error, with blue indicating low error and red indicating high error. The black arrows indicate the direction of decreasing error (i.e., the negative of the error's gradient). The white lines are iso-error contours indicating (k_L, k_G) pairs leading to identical average error. For Fig 1(a) the subwindow size is 1x1 so the local ranking step has no effect since there is only a single value to rank. The (k_L, k_G) values corresponding to the lowest error form a vertical "valley" (dark blue) around k_G = 95% between the two iso-error contours defined by the average error rates of MaxRGB (k_G = 100%) and Grayworld (approximately k_G = 80%).

For the case of 3x3 subwindows shown in Fig 1(b), the "valley" (k_L = k_G = 95%) shrinks to a "dimple" and is shifted upwards. The error also drops to 5.3 from 5.6. The "hill" on the right side of the "valley" is at k_G = k_L = 100%, which corresponds to MaxRGB without preprocessing. Increasing the subwindow size to 9x9, the "dimple" (Fig 1(c)) pivots to the left and stretches out to a "valley" again, but now a horizontal one. The minimum error now lies around k_L = 92%, k_G = 100%. Finally, for subwindows of size 21x21 the "valley" (Fig 1(d)) stretches further horizontally, and the minimum error is found at k_L = 80%, k_G = 98%.

Plots of the same kind as in Fig. 1 but based on the 568, 321 and HDR datasets instead of the 11,346 set consistently show a similar trend so they are not included here. Using grid search, the minimum error for each setting is determined. Fig. 2 plots the locations of the smallest median angular error obtained for each dataset as a function of k_L, k_G, and subwindow size.

Fig. 1 shows a remarkable "hill-valley-hill" pattern, regardless of subwindow size. It is particularly interesting that the valley is always located between the Grayworld and MaxRGB iso-error contours. This implies that there is an opportunity to improve upon the Grayworld and MaxRGB methods using the rank-based framework. For relatively large values of k_L and k_G, (e.g., 90%~99%) the resulting error is always less than both MaxRGB and Grayworld. In this sense, the rank-based method is again analogous to Shades-of-Gray [6] in that it also outperforms both Grayworld and MaxRGB for a Minkowski norm between 1 and infinity. For large values of k_L and k_G the rank method chooses high values and in this sense begins to select 'bright' pixels as suggested by Vaezi Joze et al. [10], although in that work the 'bright' values are based on R+G+B, not R, G, B separately. Also the bright pixel method uses all pixels above a certain percentile, while the rank method uses only the single pixel value at the given percentile.

The curve-shaped locus of minima in Fig. 2 indicates that the minimum angular error decreases with increasing subwindow size.

Table 2. Performance comparison on the SFU 321, grayball 11346, HDR 105 and re-processed colorchecker 568 datasets. The methods compared are Do-Nothing, Grayworld [3], Shades-of-Gray [6], Edge-Based [4] (first and second order), MaxRGB [2], MaxRGB+[1], TPS Thin-Plate Spline [8], CbyC as reported in [18], Gamut Mapping [19], N-jet [18], Bayes-GT[9], and Rank-based.

	Angular Error on Test Datasets											
	Barnard 321			HDR 105			Ciurea 11346			Gehler 568		
	Median	Trimean	Max(25%)	Median	Trimean	Max(25%)	Median	Trimean	Max(25%)	Median	Trimean	Max(25%)
Do-nothing	15	16	31	4.3	4.9	16	6.7	7.6	19	4.8	7.8	24
MaxRGB	6.4	7.8	21	4.2	5.3	15	6.0	6.8	17	9.1	9.5	21
MaxRGB+	3.0	4.4	15	3.8	5.1	15	5.3	6.0	15	4.2	6.2	18
Grayworld	7.0	7.7	23	7.3	7.7	15	6.3	6.7	14	3.7	4.1	11
SoG (p=6)	4.0	4.9	16	4.3	4.1	14	5.8	6.1	13	4.5	5.5	15
Edge-Based (1st order)	3.6	4.6	15	3.7	4.6	14	5.5	6.1	15	3.8	5.4	16
Edge-Based (2nd order)	4.5	5.6	15	4.0	4.9	14	5.4	6.2	16	4.4	5.9	17
TPS (3-fold)	1.2	0.9	3.5	--	--	--	--	--	--	2.4	2.6	6.8
CbyC (Gijsenij et al.)	6.8 [18]	--	--	5.9	--	--	6.5	7.4	17	--	--	--
Gamut Mapping	3.1 [18]	--	--	--	--	--	4.8	5.2	12	4.3	5.3	15
N-jet (complete 1-jet)	2.1 [18]	--	--	--	--	--	5.5	5.8	13	4.2	5.1	14
Bayes-GT	-	--	--	5.9	7.0	18	-	-	-	5.8	6.2	15
Rank-based (3-fold)	2.7	3.7	14	3.7	4.5	14	5.1	5.7	13	2.5	3.0	8.8

For instance, the minimum error for subwindows that cover 0.01% of the image is always larger than for subwindows that cover 1%. Based on these observations, a rule-of-thumb guideline for the choice of parameters is that the subwindow size should be relatively large (~1%) and the k_L and k_G percentile ranks should be high. The large subwindow size and high k_L together mean that the effect of the first step of the rank-based method performs a dilation that spreads out each high value to its surrounding area.

Table 1 shows that in comparison to several other illumination estimation methods the rank-based method is very competitive based on the reported median, trimean, and maximum (average of top 25%) angular errors. Since the choice of parameters is affected by the characteristics of the imaging system, the median error reported in Table 2 for the rank-based method is based on 3-fold cross validation. The comparison on the 321 set is complicated by the fact that three of the other methods are tested on a reduced set of 290 of the 321 images (namely, CbyC, Gamut Mapping and N-jet in [18]).

Conclusion

The proposed rank-based illumination estimation method is conceptually simple and testing shows it to be very effective. It uses the ranking of the RGB digital counts channel-by-channel within subwindows to produce estimates of the illumination locally. By adjusting the ranking percentile, the method varies from a median-based variant of Grayworld to MaxRGB. The analysis shows that the optimal choice of parameters lies between these two extremes. The global estimate is computed based on the rankings of the subwindow estimates. When the ranking percentile is chosen as a high number, the method tends to favour 'bright' pixels, which has been shown by Vaezi Joze et al. [10] and by by Tominaga et al. [11] to be a good strategy. Despite the rank-based method's simplicity, its performance is found to be competitive with other state-of-the art methods for estimating the chromaticity of the overall scene illumination.

References

[1] B. Funt and L. Shi, "MaxRGB Reconsidered," Journal of Imaging Science and Technology (JIST), Vol. 56, 2 (2012).

[2] E. Edwin and J. McCann, "Lightness and Retinex Theory," Journal of the Optical Society of America, 61, 1 (1971).

[3] G. Buchsbaum, "A spatial processor model for object colour perception," J. Franklin Inst. 310, 1 (1980).

[4] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," IEEE Trans. Image Process. 16, 2207 (2214).

[5] A. Gijsenij, T. Gevers, and J. van deWeijer, "Generalized gamut mapping using image derivative structures for color constancy," Int. J. Comput. Vis. 86, (2010).

[6] G. Finlayson and E. Trezzi, Shades of gray and colour constancy, Proc. IS&T/SID 12th Color Imaging Conf. pg.37–41. (2004).

[7] G.D. Finlayson, S.D. Hordley, P.M. Hubel, "Color by correlation: a simple, unifying framework for color constancy," Pattern Analysis and Machine Intelligence, IEEE Transactions on , 23, 11 (2001).

[8] L. Shi, W. Xiong and B. Funt, "Illumination Estimation via Thin-Plate Spline Interpolation," Journal of the Optical Society of America A, Vol. 28, 5, (2011).

[9] P. V. Gehler, C. Rother, A. Blake, T. P. Minka, T. Sharp, "Bayesian color constancy revisited," IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), pg. 24-26 (2008).

[10] H. R. Vaezi Joze, M. S. Drew, Graham D. Finlayson and Perla Aurora Troncoso Rey, "The Role of Bright Pixels in Illumination Estimation," 20th Color and Imaging Conference, pg. 41-46 (2012).

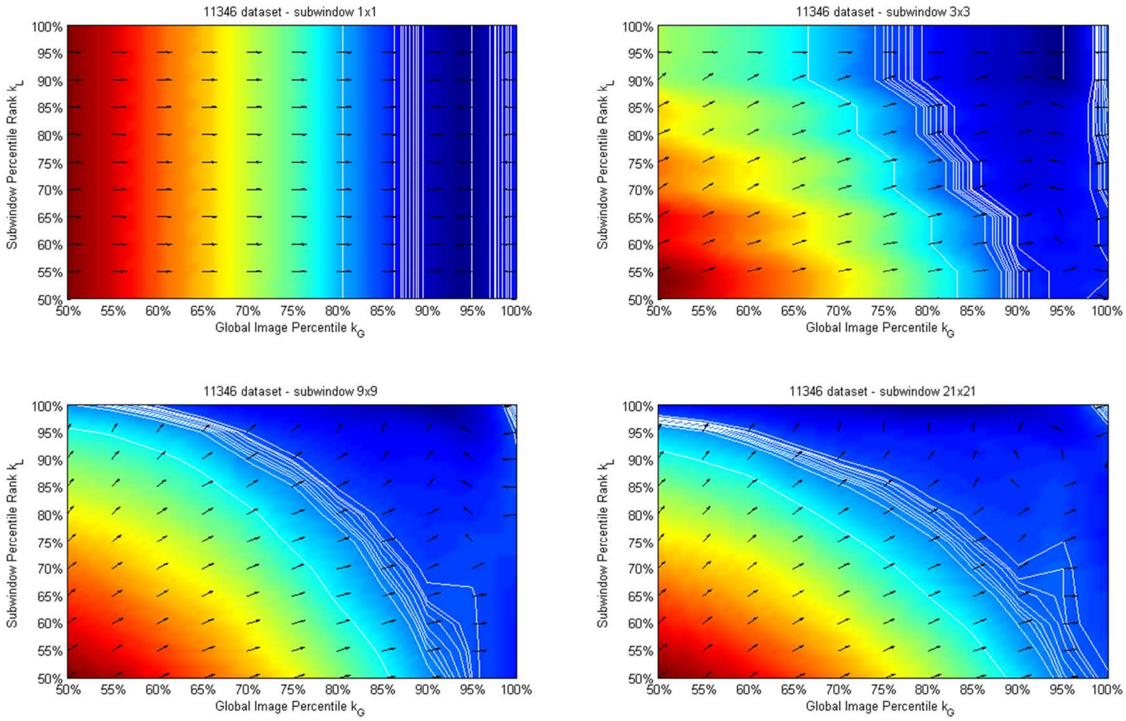

Fig 1. Four maps of the median angular illumination estimation error for different subwindow sizes on the 11346 dataset. The hotter color (red) indicates high error, and cooler color (blue) indicates low error. The thin light-coloured lines indicate iso-error contours. The arrows indicate the direction of decreasing error. The x-axis is the global ranking parameter k_G, and the y-axis is the local ranking parameter k_L.

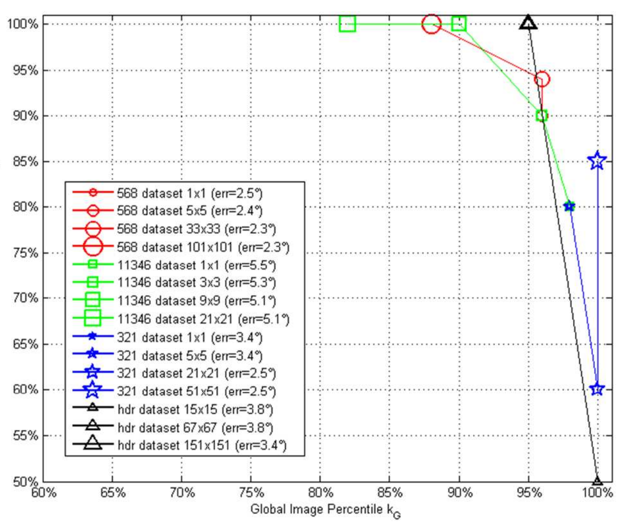

Fig 2. Plots of the (k_G,k_L) locations of the smallest median angular error for each dataset for different subwindow sizes. The legend specifies the subwindow size and the median error for the indicated point's parameters of subwindow size, k_G, and k_L.

[11] S. Tominaga, S. Ebisui, and B. Wandell, "Scene illuminant classification: brighter is better," J. Opt. Soc. Am. A 18, 55-64 (2001).

[12] A. Chakrabarti and K. Hirakawa, and T. Zickler, "Color Constancy with Spatio-Spectral Statistics," IEEE Transactions on Pattern Analysis and Machine Intelligence, 34, 8 (2012).

[13] Thomas S. Huang, Two-Dimensional Signal Processing II: Transforms and Median Filters (Berlin: Springer-Verlag, 1981) pg. 209-211.

[14] B. Weiss, "Fast median and bilateral filtering," ACM Transactions on Graphics (TOG), 25, 3 (2006).

[15] K. Barnard, L. Martin, A. Coath, and B. Funt, "A comparison of computational color constancy algorithms—Part II: Experiments with images," IEEE Trans. Image Process. 11, (2002).

[16] V.C. Cardei, B. Funt, and K. Barnard, "Estimating the scene illumination chromaticity by using a neural network," J. Opt. Soc. Am. A, 19 (2002).

[17] F. Ciurea and B. Funt, A Large Image Database for Color Constancy Research, in Proc. IS&T.SID Eleventh Color Imaging Conference, Society for Imaging Science and Technology, pg. 160-163. (2003).

[18] A. Gijsenij, T. Gevers, and J. van deWeijer, "Generalized Gamut Mapping using Image Derivative Structures for Color Constancy," Int. J. Comput. Vis., 86, 2/3 (2010).

[19] G. Finlayson, S. Hordley, and I. Tastl, "Gamut constrained illumination estimation," Int. J. Comput. Vis. 67, 93 (109).