Shi, L., Funt, B., Xiong, W., Kim, S., Kang, B., and Lee, S.D., "Illumination Estimation via Non-negative Matrix Factorization," Proc. AIC 2007 Color for Science and Industry, Midterm Meeting of the International Color Association, Hangzhou, July 2007.

Illumination Estimation via Non-Negative Matrix Factorization

L. Shi [a], B. Funt [*a], W. Xiong [a], S. S. Kim [b], B. H. Kang [b], S. D. Lee [b] and C. Y. Kim [b]

[a] Simon Fraser University, Vancouver, Canada;
[b] Samsung Advanced Institute of Technology, Korea

ABSTRACT

The problem of illumination estimation for colour constancy and automatic white balancing of digital color imagery can be viewed as the separation of the image into illumination and reflectance components. We propose using nonnegative matrix factorization with sparseness constraints (NMFsc) to separate the components. Since illumination and reflectance are combined multiplicatively, the first step is to move to the logarithm domain so that the components are additive. The image data is then organized as a matrix to be factored into nonnegative components. Sparseness constraints imposed on the resulting factors help distinguish illumination from reflectance. Experiments on a large set of real images demonstrate accuracy that is competitive with other illumination-estimation algorithms. One advantage of the NMFsc approach is that, unlike statistics- or learning-based approaches, it requires no calibration or training.

Keywords: Color Constancy, Non-Negative Matrix Factorization, Automatic White Balancing

1. INTRODUCTION

A new approach to illumination estimation for color constancy and automatic white balancing is presented based on the technique of nonnegative matrix factorization with sparseness constraints (NMFsc). In essence, the logarithm of the input color image is viewed as a matrix to be factored into independent components. The resulting components represent the scene's illumination and the reflectance. The nonnegative constraint on the factorization is important because illumination and reflectance are both nonnegative physical quantities. The sparseness constraints---illumination is non-sparse, reflectance is sparse---guide the factorization to obtain an illumination component that is relatively constant across the scene, while allowing the reflectance component to vary. Experiments on a large data set of real images show that both methods are competitive with existing illumination estimation methods.

One advantage of the NMFsc illumination method is that like a few other methods[1-4], it avoids the training step required by the many methods that rely on image statistics[5-9] or finite-dimensional models of spectra[10].

For a particular pixel in a color image, the RGB sensor response is defined by the model in Equation (1). Let $E(\lambda)$ and $S(\lambda)$ be the illumination spectral power distribution and matte surface reflectance function respectively, let $R_k(\lambda)$ be the sensor sensitivity function for a colour channel k, then the model can be defined as

$$p_k = \int E(\lambda)S(\lambda)R_k(\lambda) \qquad k = R, G, B. \quad (1)$$

Assuming the camera has narrowband spectral sensitivity functions that can be modelled by a Dirac delta function, Equation (1) simplifies to:

$$p_k = E(\lambda_k)S(\lambda_k) \qquad k = R, G, B. \quad (2)$$

By taking logarithm on both sides of the Equation (2), we have

$$\log(p_k) = \log[E(\lambda_k)] + \log[S(\lambda_k)], \ k = R, G, B. \quad (3)$$

This has the advantage that the non-linear multiplicative combination of the illumination and reflectance becomes linear.

For an image or image subwindow arranged as a vector, Equation (2) yields

$$\mathbf{I} = \mathbf{E} \circ \mathbf{S}, \quad (4)$$

where **I** is a 2D image, and **E** and **S** are the illumination and surface reflectance images, respectively. The operator $\circ$ denotes element-wise multiplication. Applying logarithms again, we have

$$\log \mathbf{I} = \log \mathbf{E} + \log \mathbf{S}. \quad (5)$$

Here, log **E** is the illumination term, and log **S** is the reflectance term. They correspond to the "illumination image" and "reflectance image" in log space.

[*] Further author information: (Send correspondence to B. Funt)
B. Funt: E-mail: funt@sfu.ca. Telephone: 1 604 291 3126 Website: www.cs.sfu.ca/~funt

Generally, illumination is relatively constant across an image, while the reflectance varies. The reflectance image in log space can be further decomposed and represented as a weighted linear combination of "feature" reflectances.

$$\log \mathbf{S} = \sum_{i=1}^{M} \mathbf{F}_i h_i \,, \tag{6}$$

where $\mathbf{F}_i$ are independent reflectance features and h_i are the weighting coefficients. In order to be independent, these features need to be "non-overlapping," which means that most entries of the vectors are zeros, and the non-zero entries appear at distinct locations. These non-overlapping, sparse features can be thought of as building blocks from which the image is constructed. Therefore, in log space, by Equation (5) and (6), the image can be represented in terms of the illumination and M surface features as

$$\log \mathbf{I} = \log \mathbf{E} + \sum_{i=1}^{M} \mathbf{F}_i h_i \tag{7}$$

Since we expect the illumination to vary slowly across an image, log **E** should be a non-sparse vector. On the other hand, the reflectance term log **S** should be a sparse vector. Ideally, the feature vectors $\mathbf{F}_i$ should be sparse enough that there is no overlap between them so that they are completely independent features.

2. ESTIMATING ILLUMINATION USING NMFSC

Non-negative matrix factorization creates a non-negative approximation for a given set of input data that represents the data in terms of a linear combination of non-negative basis features[11]. In the context of color imagery, we will use it to represent the log image data in terms of a linear combination of log illumination and log reflectance.

Let us assume that the data consists of T measurements of N non-negative scalar variables. Denoting the (N-dimensional) measurement vectors by $\mathbf{v}^t$ ($t = 1, \ldots, T$), a linear approximation of each data vector is given by

$$\mathbf{v}^t \approx \sum_{i=1}^{M} \mathbf{w}_i h_i = \mathbf{W}\mathbf{h}^t \,, \tag{8}$$

where $\mathbf{W}$ is an $N \times M$ matrix containing the basis vectors $\mathbf{w}_i$ as its columns, and $\mathbf{h}^t$ is the vector of coefficients h_i. Arranging vectors $\mathbf{v}^t$ as columns of an $N \times T$ matrix $\mathbf{V}$, we have

$$\mathbf{V} \approx \mathbf{W}\mathbf{H} \,, \tag{9}$$

where each column of **H** contains the coefficient vector $\mathbf{h}^t$ corresponding to the measurement vector $\mathbf{v}^t$. Written in this form, it becomes apparent that this linear data representation is simply a factorization of the data matrix. Principal component analysis, independent component analysis, vector quantization, and non-negative matrix factorization can all be viewed as matrix factorization methods, with different choices of objective functions or constraints. Whereas PCA and ICA do not restrict the signs of the entries of **W** and **H**, NMF requires all entries of both matrices to be non-negative, which means that the data is described in terms of additive components only.

The concept of 'sparse coding' refers to a representational scheme where only a few units are used to represent typical data vectors[12]. In effect, this implies that the majority of units take values close to zero, with only a few having significantly non-zero values.

Hoyer[12] adopts a sparseness measure based on the relationship between the *L1* norm and the *L2* norm defined as

$$s(\mathbf{x}) = \frac{\sqrt{n} - \frac{\sum |x_i|}{\sqrt{\sum x_i^2}}}{\sqrt{n} - 1} \,, \tag{10}$$

where N is the dimensionality of x. $\sum |x_i|$ is the *L1* norm, and $\sqrt{\sum x_i^2}$ is the *L2* norm. This function evaluates to unity if and only if **x** contains a single non-zero component, and takes a value of zero if and only if all components are equal. It also interpolates smoothly between the two extremes.

Generally, an image will contain multiple surface reflectance features, so when subwindow sample blocks are drawn from the image, each block should contain some subset of those features. NMFsc provides a way to identify a set of basis vectors to represent these surface reflectance features plus a single illumination feature. Since each subwindow is described by using strictly additive positive components, it is a linear combination of those feature vectors.

The imaging model in Equation (7) and NMFsc in Equation (8) have parallel structure, so that the imaging model can be reformatted in terms of an NMF approximation:

$$\begin{aligned} \mathbf{v}^t \approx \sum_{i=0}^{M} \mathbf{w}_i h_i &= \mathbf{w}_0 h_0 + \sum_{i=1}^{M} \mathbf{w}_i h_i \\ &= \log \mathbf{E} + \sum_{i=1}^{M} \mathbf{F}_i h_i \end{aligned} \tag{11}$$

In this case, $\mathbf{v}^t$ corresponds to $\log \mathbf{I}$ in Equation (7) and represents the data from one of the image blocks. Since $\mathbf{w}_0 h_0$ takes the role of $\log \mathbf{E}$, the basis vector $\mathbf{w}_0$ is the "illumination" basis with weighting factor h_0. M represents the number of features present in the data. Since

$\sum_{i=1}^{M} \mathbf{w}_i h_i$ takes the role of $\sum_{i=1}^{M} \mathbf{F}_i h_i$, the basis vectors $\mathbf{w}_i$ are the feature reflectance basis vectors with weighting factors h_i. The weighting factors determine how strongly the corresponding feature reflectances should appear in this image block. For instance, if h_i equals to zero, it means the feature is absent from this block; if h_i is large, it means the feature is strongly visible in this sample block.

By taking T sample sub-windows from the image and constructing the data matrix **V**, where the log RGB channels of each block are appended and stored as a column, we can then use NMFsc to solve for the basis matrix, **W,** and thereby obtain the illumination basis vector and the feature reflectance basis vectors. In other words, NMFsc decomposes **V** into the illumination and reflectance components that are the key to color constancy and automatic white balancing.

Equation (7) is a purely additive model, which means NMF is an appropriate approach to solving for the basis. All basis vectors, including the feature reflectance images (the "building blocks"), along with the illumination image, are required to be non-negative. This requires the input data matrix **V** to be non-negative too. The model is applied to the logarithm of the original image data, so there is the possibility of both positive and negative values. Simply scaling the original image data to (0,1] ensures that all pixel values in log space will be negative or zero. Since the coefficient matrix **H** is always non-negative, we negate both **V** and **W** to make everything completely non-negative.

NMFsc allows the sparseness for each basis vector to be controlled individually. In our model, the illumination basis vector is supposed to be non-sparse, making its components relatively similar, while the reflectance basis vectors are supposed to be sparse. In addition, with NMFsc the sparseness of each portion of a single basis vector can be controlled separately. This feature is important because in the formulation the RGB components needed to be packed into one vector. If a small sparseness value is set for the vector as a whole then the illumination basis will be similar across all the RGB channels, collectively leading to grey as the illumination estimate. To avoid this problem, the sparseness of the illumination basis vector needs to be controlled individually for the R, G, and B segments of the vector. In other words, the illumination vector must be divided into three segments and the same sparseness applied to each. For the reflectance basis vectors, a very sparse vector means that most of the entries are zeros. This property of high sparseness allows the reflectance basis vectors to be orthogonal and independent.

Hence, NMFsc is an approach for solving the illumination-reflectance model globally, in that the factorization aims to minimize the objective functions based on the data matrix that includes all three channels. This is an advantage over those methods that estimate the illumination and reflectance for each colour component independently.

The proposed algorithm based on Equation (9) using the NMFsc approach is:

1. Scale the input image values to (0,1]
2. Take N sample blocks from the image
3. Take the logarithm of the RGB values in these blocks.
4. For the data from each block, concatenate the color channels into a vector.
5. Suppose there are M different surfaces appearing in N blocks (M<N)
   - 5.1. Apply NMFsc to find M+1 basis vectors
   - 5.2. Set the sparseness constraint of the 1st basis close to 0 since it represents the illumination
   - 5.3. Set sparseness constraints of the 2nd to (M+1)th bases close to 1 since they represent the surface features
6. Antilog the illumination basis
7. The average R, G, B from the channels of the antilog of the illumination basis yields the RGB color of the scene illumination.

The parameter M represents the number of feature reflectances assumed to be present in the input image; however, the correct value of M is unknown and could differ from image to image. Experimentally, we found that fixing M at 5 for all images worked well.

In the above development, an image was assume to contain multiple reflectance features. An image contains M feature reflectances with at least one feature appearing in each image subwindow. Data was collected from multiple subwindows to form the data matrix for NMFsc. However, instead of M reflectance features, suppose that we describe the scene as a single more complex reflectance feature under a single illumination and apply NMFsc. In this case, there is only one subwindow—the entire image—and there will be only a single reflectance basis vector.

Equation (9) with $M = 1$ combined with Equation (6) becomes

$$\mathbf{v}^t \approx \sum_{i=0}^{1} \mathbf{w}_i h_i = \mathbf{w}_0 h_0 + \mathbf{w}_1 h_1 = \log \mathbf{E} + \log \mathbf{S} \qquad (12)$$

Here again, $\mathbf{w}_0 h_0 = \log \mathbf{E}$ so the basis vector $\mathbf{w}_0$ is the "illumination" basis with weighting factor

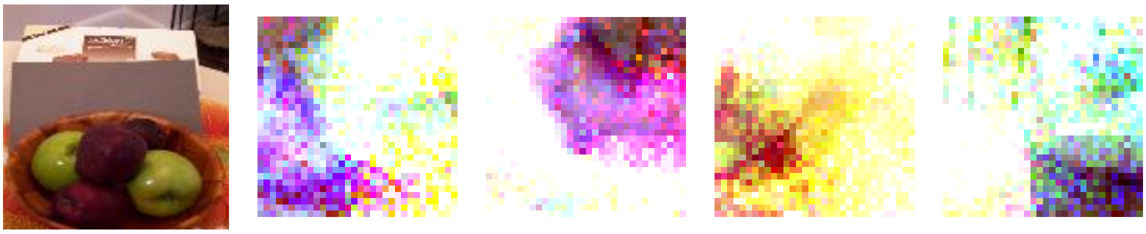
h_0. Similarly, $\mathbf{w}_1 h_1 = \log \mathbf{S}$ so the basis vector $\mathbf{w}_1$ is the feature reflectance basis vector with weighting factor h_1. The goal in Equation (12) is to split the input color image into an illumination component and reflectance component in log space. How NMFsc does the split depends on the choice of sparseness constraints for the two components.

Since NMFsc should return two basis vectors $\mathbf{w}_1$ and $\mathbf{w}_2$, it requires an input data matrix of at least two columns. Rather than taking two distinct sample subwindows as the input data, we construct the data matrix $\mathbf{V}$ with two identical columns. Each column is a vectorized version of the full input image. It is no longer necessary to estimate the parameter M because in this case it always equals 1. Note also, that whereas the location of each pixel matters in the multiple-reflectance model, location has no effect in the single-reflectance model.

3. EXPERIMENTS

The NMFsc illumination estimation method is evaluated on a number of different image databases. Both the multiple-reflectance-feature reflectance and single-reflectance-feature approaches are tested.

The first set of tests is with multiple features. Figure 1 gives an example.



(a) (b) (c) (d) (e)

Figure 1. The reflectance basis vectors (contrast enhanced for visualization) based on the multiple-feature reflectance model: (a) 128x128 input image ; (b)-(e) are the reflectances basis vectors F_i using 32x32 subwindows.

Method	Angular Degrees			Distance($\times 10^2$)		
	Mean	RMS	Max	Mean	RMS	Max
GW	7.69	9.38	42.28	5.97	7.47	38.33
SoG	7.50	8.93	34.52	5.50	6.57	27.67
MAX RGB	9.99	11.76	27.42	7.24	8.60	21.72
NMFsc (M = 5)	7.66	8.96	34.79	5.59	6.57	26.99
NMFsc (M = 1)	6.82	8.15	38.27	5.11	6.18	32.74

Table 1. Comparison of NMFsc to SoG, Max RGB, Grayworld performance. The results involve testing on the large natural image dataset, with no real-data training required. Errors are reported in terms of both the RMS angular chromaticity and distance error measures.

NMFsc is applied with $M = 4$, sparseness of the illumination basis is set to 0.005 for the R, G, and B channels separately, and the sparseness of the feature reflectance basis is set to be 0.45. Figure 1 (b)-(e) shows the feature reflectance basis vectors (i.e., the antilog of the $\boldsymbol{w}_i$'s in Equation (11) with $1 \le i \le M$).

The second test provides statistical results about the accuracy of NMFsc-based illumination estimation. The test set is extracted from the large dataset of natural images representing a variety of indoor and outdoor scenes under different light conditions that Ciurea et. al.[13] measured with a grayball attached to a digital video camera. The original image database includes 11,346 images. However, many of these images have very good color balance (i.e., RGB of the gray ball is gray) which could bias the testing of the illumination-estimation methods. Therefore, we eliminated from the data set the majority of the correctly balanced images so that the overall distribution of the illumination color is more uniform. The resulting data set contains 7661 images. The grayball appears in the lower right-hand quadrant of every original image, so for testing that quadrant is cropped from every image.

The 7,661 images are tested based on the SoG, Max RGB, and Grayworld methods, as well as both our multiple-reflectance and single-reflectance methods. The accuracy of various illumination estimation methods (Shades of Gray, Max RGB, Grayworld, single-reflectance NMFsc, and multiple-reflectance NMFsc) applied to the 7,661 images is listed in Table 1. In the case of the multiple-reflectance based estimation, each image is resized to 64x64 pixels, and divided into sixteen 16x16 subwindows. The number of reflectance features *M* is set to be 5; the sparseness of the illumination and the reflectance bases are 0.001 and 0.45, respectively. The average computation time for processing one image is 0.83 seconds. In the case of the single-reflectance based estimation, each image is also resized to 64x64. The sparseness of the illumination and the reflectance bases are 0.001 and 0.45, respectively. With $M = 1$, the average computational time for processing one 64x64 image is 2.43 seconds.

4. CONCLUSION

The experiments show that nonnegative matrix factorization with sparseness constraints provides a method of separating a color image into its illumination and reflectance components. The accuracy of the NMFsc method is competitive with other illumination-estimation algorithms. One possible disadvantage of the approach is that existing factorization algorithms are iterative, and in comparison to some of the other illumination-estimation algorithms, somewhat costly in terms of computation. A particularly good feature of the NMFsc approach is that it requires no training.

REFERENCES

1. G. D. Finlayson and E. Trezzi, "Shades of Gray and Colour Constancy," *Proc. IS&T/SID Twelfth Color Imaging Conference: Color Science, Systems and Applications*, *Society for Imaging Science and Technology*, Scottsdale, 37-41, Nov. 2004.
2. B. Funt, F. Ciurea, and J. McCann "Retinex in Matlab," *Proceedings of the IS&T/SID Eighth Color Imaging Conference: Color Science, Systems and Applications*, 112-121, 2000.
3. G.D. Finlayson. "Retinex viewed as a gamut mapping theory of color constancy," *Proc. AIC International Color Association Color* 97 Vol. 2, Kyoto, Japan, 527-530, 1997.
4. B. Buchsbaum. "A spatial processor model for object color perception," *Journal of the Franklin Institute*, Vol. 31, 1-26, 1980.
5. B.V. Funt, and W.H. Xiong. "Estimating Illumination Chromaticity via Support Vector Regression," *Proceedings of the IS&T/SID Twelfth Color Imaging Conference: Color Science, Systems and Applications*, 47-52, 2004.
6. V. Cardei, B. Funt. and K. Barnard, "Estimating the Scene Illumination Chromaticity Using a Neural Network," *Journal of the Optical Society of America A*, Vol. 19, No. 12, Dec. 2002.
7. G. D. Finlayson, S. D. Hordley and P. M. Hubel, "Color by Correlation: A Simple, Unifying Framework for Color Constancy," *IEEE Trans. on Pattern Analysis And Machine Intelligence* Vol.23, No. Nov. 2001.
8. G. D. Finlayson, S. D. Hordley, P. M. Morovic, "Colour Constancy Using the Chromagenic Constraint," *Proc. Comp. Vision and Pattern Recognition.* 1079-1086, 2005.
9. C. Rosenberg, M. Hebert, and S. Thrun. "Image Color Constancy Using KL-Divergence". In *IEEE International Conference on Computer Vision*, 2001.
10. L. T. Maloney, and B. A. Wandell, "Color Constancy: A Method for Recovering Surface Spectral Reflectance," *Journal of the Optical Society of America A*, Vol. 3, No. 1, 29-33, 1986.
11. D. D. Lee and H. S. Seung. "Algorithms for non-negative matrix factorization". *In Advances in Neural Information Processing Systems* 13, 556-562, 2001
12. P.O. Hoyer, "Non-negative Matrix Factorization with sparseness constraints," *Journal of Machine Learning Research*, Vol. 5, 1457-1469, 2004.
13. F. Ciurea and B. Funt, A Large Image Database for Color Constancy Research, *Proc. IS&T.SID Eleventh Color Imaging Conference, Society for Imaging Science and Technology*, 160-163, Nov. 2003.
14. A. Pascual-Montano, J.M. Carazo, K. Kochi, D. Lehmann, R. D. Pascual-Marqui, "Nonsmooth Nonnegative Matrix Factorization (nsNMF)," *IEEE Trans, on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 3, 403-415, 2006.