

Joint Image GMM and Shading MAP Estimation

Alexander Shekhovtsov and Václav Hlaváč

Faculty of Electrical Engineering, Czech Republic, <http://cmp.felk.cvut.cz>
shekhovt@cmp.felk.cvut.cz hlavac@fel.cvut.cz

Abstract

We consider a simple statistical model of the image, in which the image is represented as a sum of two parts: one part is explained by an i.i.d. color Gaussian mixture and the other part by a (piecewise-) smooth grayscale shading function. The smoothness is ensured by a quadratic (Tikhonov) or total variation regularization. We derive an EM algorithm to estimate simultaneously the parameters of the mixture model and the shading. Our algorithms for both kinds of the regularization solve for shading and mean parameters of the mixture model jointly.

I. Introduction

This work proposes a novel model to estimate shading in natural images¹. We assume some areas of the image are brighter or darker because of the illumination/shadows. We refer to this effect as shading. The shading must be simple a priori – either smooth or piecewise-smooth function over the image domain. The image with the shading removed is then explained by a Gaussian mixture model (GMM) with unknown parameters. We pose the estimation problem as maximizing the image likelihood jointly in shading and GMM parameters. This maximum a posteriori problem (MAP) corresponds to the optimization problem, where smoothness of the shading is imposed via regularization.

Contribution. We showed that in the considered model it is possible to optimize the likelihood w.r.t. shading and mean parameters of the GMM jointly, which makes the algorithms less prone to getting stuck in local minima. We derived a simple algorithm for the case of the quadratic regularization on the shading, which monotonously improves the objective. For the case of Total Variation (TV) regularization, we derived algorithms which follow existing TV methods,

but estimate GMM mean parameters jointly with the shading and put these algorithms into the EM loop.

Related Work. This work was mostly inspired by [1], where the segmentation model includes segmentation and shading. In the case of MAP shading recovery, the estimation of the shading (by a discrete optimization) is alternated with estimation of all appearance parameters, which is inferior to our method.

A lot of previous work, e.g. [2], [3], exploit illumination-invariant characteristics of the color image such as chromaticity. The shading could be recovered from edges present in the image but absent in the illumination-invariant representation. Such edges are considered to be produced by shadows and they are integrated to reconstruct the shading. In contrast, our model can estimate soft shadows and remains meaningful also for grayscale images. A variational formulation of the problem was proposed in [4]. However, it seems to be too simplistic, and the results shown in further work [5], [6] are not convincing. Our problem of decomposing the image into shading and a general color mixture components is closely related to several other decompositions. Image denoising can be viewed as decomposition into signal and noise parts. A Gaussian noise model would correspond to the special case of a single-component GMM. Structure-texture decomposition [7], [8] seeks for piecewise smooth structure and possibly largely oscillating texture components. For grayscale images, the shading in our model corresponds to the structure in [8] (and the same regularization is imposed). The GMM log likelihood term can be viewed as a parametrized regularization on the texture component.

A related problem, inverse lighting, exploits additional a priori knowledge about scene geometry and albedo.

II. Model

Notation. Let $\Omega \subset \mathbb{R}^2$ be a continuous image domain. For a function $u: \Omega \rightarrow \mathbb{R}^p$, its value at a point $s \in \Omega$ is denoted as u_s and belongs to \mathbb{R}^p . Euclidean

¹The work was supported by EU projects FP7-ICT-247870 NIFTi and FP7-ICT-247525 HUMAVIPS and CR project 1M0567 CAK.

norm is denoted by $\|\cdot\|$. The point-wise norm of u , denoted as $|u|: \Omega \rightarrow \mathbb{R}$, is defined by $|u|_s = \|u_s\|$.

Let $I: \Omega \rightarrow \mathbb{R}^3$ be the RGB color image. Let $h: \Omega \rightarrow \mathbb{R}$, satisfying

$$\int_{\Omega} h_s ds = 0, \quad (1)$$

be unknown shading (lighting) for the image I . We define the likelihood of observing the image I given the shading h as

$$p(I|h; \theta) = \exp \left\{ \int_{\Omega} \log p(I_s|h_s; \theta) ds \right\}, \quad (2)$$

where $p(I_s|h_s; \theta)$ is the likelihood of observing color I_s given the shading h_s in the point s and θ is a vector of parameters. In the case of discrete domain and pixel colors i.i.d. from $p(I_s|h_s; \theta)$, their joint probability would factor as

$$\prod_s p(I_s|h_s; \theta) \quad \left(= \exp \sum_s \log p(I_s|h_s; \theta) \right). \quad (3)$$

The model (2) is a continuous analogue of (3).

In the additive shading model, we define “unshaded” image as $U(h)_s = I_s - h_s \mathbf{1}_3$, where $\mathbf{1}_3 = (1 \ 1 \ 1)^T$. Hence the shading h is the amount of white light added to the unshaded image to produce the observed image. Whereas this formula does not correspond to a physical law, it should be understood as a local approximation to the effect we want to model.

The unshaded image color at a point s is explained by a Gaussian mixture model:

$$p(I_s|h_s; \theta) = \sum_k \pi_k \mathcal{N}(U(h)_s; \mu_k, \Sigma_k), \quad (4)$$

where

$$\mathcal{N}(x; \mu_k, \Sigma_k) = \frac{1}{(2\pi)^{3/2} \det(\Sigma_k)^{1/2}} e^{(x-\mu_k)^T \Sigma_k^{-1} (x-\mu_k)} \quad (5)$$

is a Gaussian pdf with mean $\mu_k \in \mathbb{R}^3$ and covariance $\Sigma_k \in \mathbb{R}^{3 \times 3}$ and $\pi_k \in [0, 1]$, $\sum_k \pi_k = 1$, are mixture coefficients. We let θ denote the vector of all parameters, $\theta = (\pi_k, \mu_k, \Sigma_k)_{k=1}^K$.

The main assumption about the shading is that it should be a smoothly varying function. We consider two priors (regularizations) enforcing smoothness, given by the following equation with $\rho = 1, 2$:

$$p(h) \propto \exp \left\{ -\frac{\gamma}{2} \int_{\Omega} |\nabla h|_s^\rho ds \right\} \quad (6)$$

This functional assigns low probability to the shadings which have strong gradients. The choice $\rho = 2$ allows easier optimization, whereas $\rho = 1$ (Total Variation) is more appropriate to model piecewise smooth shadings, which occur naturally in 3D scenes. Note that parameter γ has a different meaning for the two cases and it depends on the scale of h .

Our goal is to estimate shading h and the unknown parameters θ . We choose the joint MAP estimate $(h^{\text{MAP}}, \theta^{\text{MAP}})$, maximizing the joint likelihood,

$p(I, h; \theta) = p(I|h; \theta)p(h)$. Taking the logarithm, the problem can be written as maximizing $E(h, \theta) \stackrel{\text{def}}{=}$

$$\int_{\Omega} \log \sum_k \pi_k \mathcal{N}(U(h)_s; \mu_k, \Sigma_k) ds - \frac{\gamma}{2} \int_{\Omega} |\nabla h|^\rho ds. \quad (7)$$

We refer to the case $\rho = 2$ as quadratic regularization model, and case $\rho = 1$ as TV model.

III. Optimization

Quadratic Regularization. Consider maximizing (7) with $\rho=2$. The EM algorithm is obtained as following (e.g., [9]). It follows from the generalized inequality of arithmetic and geometric means that for any numbers $\alpha_{k|s} > 0$, $s \in \Omega$, $k = 1 \dots K$, such that $\sum_k \alpha_{k|s} = 1$ for all s , there holds

$$E(h, \theta) \geq \int_{\Omega} \left(\sum_k \alpha_{k|s} \log \pi_k G_{\Sigma_k}(U(h)_s - \mu_k) - \sum_k \alpha_{k|s} \log \alpha_{k|s} \right) ds - \frac{\gamma}{2} \int_{\Omega} |\nabla h|^\rho ds \stackrel{\text{def}}{=} E(h, \theta, \alpha). \quad (8)$$

When α is fixed, it is easy to differentiate and optimize $E(h, \theta, \alpha)$ in (h, θ) (M-step) and when h and θ are fixed, the inequality in (8) can be tightened by maximizing $E(h, \theta, \alpha)$ w.r.t. α (E-step). Thus we are to find $\arg \max_{h, \theta, \alpha} E(h, \theta, \alpha)$ subject to the constraints.

The E-step is the same as for the usual Gaussian mixture model:

$$\alpha_{k|s} = \frac{\pi_k \mathcal{N}(U(h)_s; \mu_k, \Sigma_k)}{\sum_{k'} \pi_{k'} \mathcal{N}(U(h)_s; \mu_{k'}, \Sigma_{k'})}. \quad (9)$$

It is seen that optimal numbers $\alpha_{k|s}$ correspond to the estimate of the conditional probability that mixture component k has generated color $U(h)_s$.

When α is fixed, optimizing in π_k subject to the constraint $\sum_k \pi_k = 1$ gives the update

$$\pi_k = \int_{\Omega} \alpha_{k|s} ds / \int_{\Omega} 1 ds, \quad (10)$$

where we used that $\sum_k \alpha_{k|s} = 1$ for all s .

We can optimize w.r.t. shading h and means μ simultaneously. The corresponding first order necessary optimality conditions are: $\nabla_h E = 0$, $\nabla_{\mu} E = 0$, where $\nabla_h E$, $\nabla_{\mu} E$ are Frechet derivatives. Standard derivation (see [10]) gives

$$0 = \sum_k \alpha_{k|s} \mathbf{1}_3^T \Sigma_k^{-1} (I_s - h_s \mathbf{1}_3 - \mu_k) + \gamma \Delta h_s \quad \forall s \quad (11a)$$

$$0_3 = \int_{\Omega} \alpha_{k|s} \Sigma_k^{-1} (I_s - h_s \mathbf{1}_3 - \mu_k) ds \quad \forall k. \quad (11b)$$

The system (11) and constraint (1) is a system of linear equations in h, μ . While the equations are differential in h , in the discretization they become just a sparse linear system and can be solved numerically with standard methods.

Optimizing w.r.t. covariance matrices Σ_k under

fixed remaining variables gives the update

$$\Sigma_k = \frac{\int_{\Omega} \alpha_k |s| (U(h)_s - \mu_k) (U(h)_s - \mu_k)^{\top} ds}{\int_{\Omega} \alpha_k |s| ds}. \quad (12)$$

To prevent appearance of degenerate components we further make sure that the smallest eigenvalue of Σ_k is not smaller than some σ_0 by projecting Σ_k on the constraint $\min \text{eig}(\Sigma_k) \geq \sigma_0$. Let $\Sigma_k = U \text{diag}(\lambda) U^{\top}$, then the projection is given by

$$\tilde{\Sigma}_k = U \text{diag}(\max(\lambda, \sigma_0)) U^{\top}. \quad (13)$$

Discretization. The discretization is built as follows. Domain Ω is quantized into 2D unit grid. Symbols $\int_{\Omega} \square ds$ are to be replaced with $\sum_{s \in \Omega} \square$. Derivatives are approximated as finite differences. For the exact choice of the approximation and boundary conditions, we followed recommendations of [11]. With variables and updates considered discretized, the algorithm **Alg1** is as follows. The input is the image I , initial GMM $(\pi_k, \mu_k, \Sigma_k \mid k = 1 \dots K)$ and initial shading $h = 0$. The algorithm iteratively updates variables $(\alpha, \pi_k, \Sigma_k)$, via (9), (10), (12), (13), and variables (h, μ) via solving linear system (11), (1). Each of the updates maximizes the objective $E(h, \theta, \alpha)$ w.r.t. the corresponding variable, and therefore is guaranteed to make a monotonous improvement to the objective.

TV Regularization. Consider maximizing (7) with $\rho=1$. We first proceed as in the previous case. We introduce expectation variables α and inequality (8) holds. Optimization w.r.t. variables α, π, Σ is performed by the update equations (9), (10), (12). We then focus on optimizing (7) w.r.t. shading h and means μ . The derivative of the TV regularization $-\int_{\Omega} |\nabla h|_s ds$ is the non-linear expression $-\nabla \cdot \frac{\nabla h}{|\nabla h|}$. It is not straightforward how to perform optimization w.r.t. h or (h, μ) jointly. We tested two algorithms described in [11]. Our derivation introduces optimization w.r.t. additional parameters μ into these algorithms. Full details are given in our technical report [10]. Here, we will review, a more accurate, dual method, which is derived following [12], [11]. The difficult term $|\nabla h|$ is represented as the maximization problem:

$$\begin{aligned} \|(\nabla h)_s\| &= \max_{u_s} u_s \cdot (\nabla h)_s. \\ \text{s.t. } \|u_s\| &\leq 1 \end{aligned} \quad (14)$$

Introducing additional unknown $u: \Omega \rightarrow \mathbb{R}^2$, the maximization of (8) can be written as

$$\max_{h, \theta, \alpha} \min_{|u| \leq 1} E(h, \theta, \alpha, u), \quad (15)$$

with

$$E(h, \theta, \alpha, u) \stackrel{\text{def}}{=} \int_{\Omega} (\dots) ds - \frac{\gamma}{2} \int_{\Omega} u_s \cdot (\nabla h)_s ds, \quad (16)$$

where the bracket (\dots) is the same as in (8). Noting that $E(h, \theta, \alpha, u)$ is convex in u , concave in (h, μ) and the constraints on u are convex, we can swap in (15)

max over (h, μ) and min over u :

$$(15) = \max_{\alpha, \pi, \Sigma} \min_{|u| \leq 1} \max_{h, \mu} E(h, (\pi, \mu, \Sigma), \alpha, u). \quad (17)$$

The inner maximization problem can be solved explicitly (see [10]) and an analytic expression can be written for the derivative of $E((\pi, \mu, \Sigma), \alpha, u)$ w.r.t. u . Then, the minimization problem in (17) can be solved by a projected gradient ascend on u (following [11]). Let us denote the corresponding algorithm for the full problem (15), using dual updates, as **Alg2D**. It iteratively updates variables $(\alpha, \pi_k, \Sigma_k)$ as (9), (10), (12) and variables (h, u, μ) by the projected grad. ascend on u .

IV. Experiments

Both considered algorithms are at best to output a local maximum, because GMM estimation is a non-convex problem by itself. Alg1 is monotonous and is guaranteed to converge. Alg2D only becomes monotonous when the inner minimization is solved to the optimality. This inner minimization however, has linear convergence and may require long time to converge to a sufficient accuracy. We got successful results with Alg2D when it was starting from a good initial estimate, *e.g.* by Alg1. For all experiments we adjusted regularization parameter γ in order to obtain visually the best results. For the setting of the other parameters see [10].

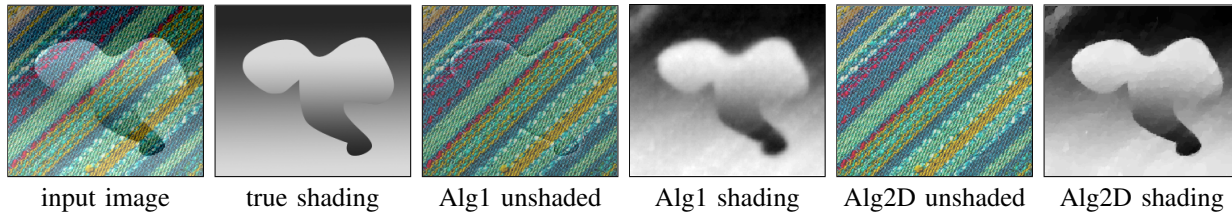
Texture unshading. As a simple test we prepared the following artificial example. We took a real world texture and added to it a known piecewise-continuous smooth shading (Fig.1 top). This experiment is similar to [1, Fig.1], only we took a texture instead of i.i.d. Gaussian noise. The colorful texture allows to test a general case where Gaussian mixture model does not degenerate to a single Gaussian. As expected, quadratic model (Alg1) gives larger error around discontinuities of the shading, and the error is removed by the TV model (Alg2).

Collor Illusion. The second experiment is on the famous color illusion picture (Fig.1 middle-left). Square B in the picture is perceived as white, however it is absolutely identical in color to square A , which is perceived as black. The challenge is whether we can undo this effect by estimating the proper shading. Inspection of the colors in the results shows that it is indeed the case.

V. Conclusion

We presented a simple model incorporating image shading with quadratic and total variation regularizations. The model with TV regularization performs better when there are sharp boundaries of the objects or sharp shadows (the shading is piecewise-smooth). However, the inner minimization problem in Alg. 2D

Synthetic shading



Color Illusion

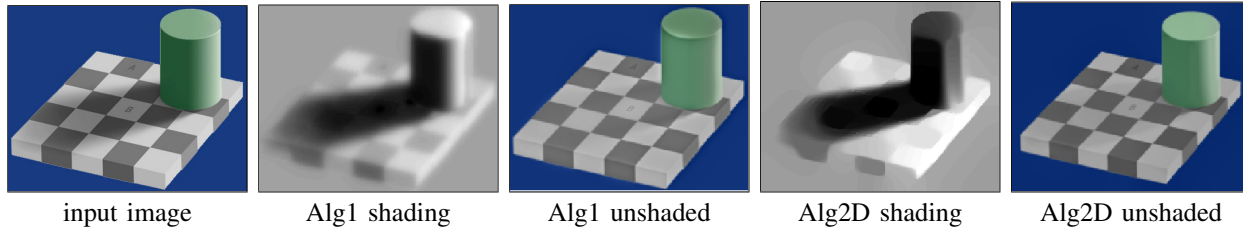


Image from [6]

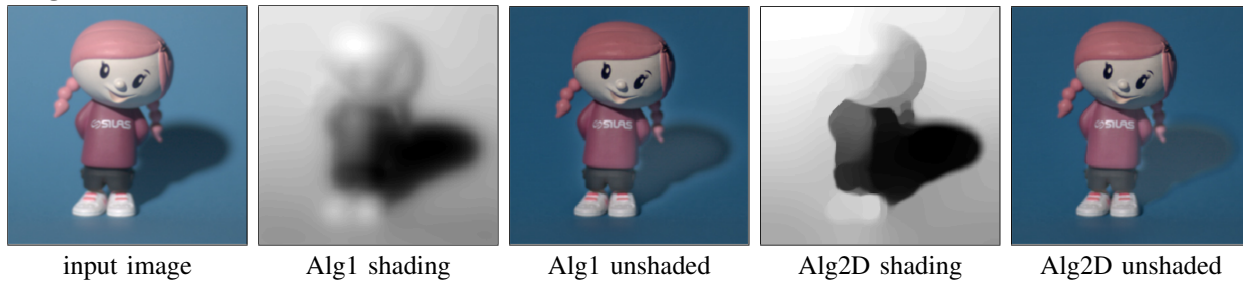


Figure 1. Experiments (see text).

may require lots of iterations to converge. Alternatively, to recover a better shading, a segmentation should be introduced into the model, similar to [1]. Different objects in the image could then have different GMM models and shadings, and the sharp boundary would be ensured by the segmentation. We expect that given the correct segmentation, shadings for the individual segments would be more accurate. For this purpose, the quadratic model may turn out more appropriate.

References

- [1] D. Schlesinger and B. Flach, "A probabilistic segmentation scheme," in *Pattern Recognition*, ser. LNCS, vol. 5096, 2008, pp. 183–192.
- [2] B. V. Funt, M. S. Drew, and M. Brockington, "Recovering shading from color images," in *ECCV*, 1992, pp. 124–132.
- [3] G. Finlayson, S. Hordley, C. Lu, and M. Drew, "On the removal of shadows from images," *PAMI*, vol. 28, no. 1, pp. 59–68, January 2006.
- [4] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Sobel, "A variational framework for retinex," *IJCV*, vol. 52, no. 1, pp. 7–23, April 2003.
- [5] M. Elad, R. Kimmel, D. Shaked, and R. Keshet, "Reduced complexity retinex algorithm via the variational approach," *JVCIR*, vol. 14, no. 4, pp. 369–388, December 2003.
- [6] H. Takahashi, T. Saito, and T. Komatsu, "Variational retinex algorithm with its application to a high-quality chroma key," in *ICIP*, 2006, pp. 977–980.
- [7] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, 2001.
- [8] J. Aujol, G. Gilboa, T. Chan, and S. Osher, "Structure-texture image decomposition: Modeling, algorithms, and parameter selection," *IJCV*, vol. 67, no. 1, pp. 111–136, April 2006.
- [9] T. Minka, "Expectation-maximization as lower bound maximization," 1998.
- [10] A. Shekhovtsov and V. Hlavac, "Joint image GMM and shading MAP estimation," Czech Technical University, Research Report, 2010, <http://cmp.felk.cvut.cz/~shekhovt/publications/shekhovtsov-10-TV-GMM-TR.pdf>.
- [11] M. Unger, T. Pock, and H. Bischof, "Interactive globally optimal image segmentation," Inst. for Computer Graphics and Vision Graz University of Technology, Technical Report ICG-TR-08/02, July 2008.
- [12] J. Carter, "Dual methods for total variation-based image restoration," Ph.D. dissertation, UCLA, Los Angeles, CA, USA, 2001.