

## Research Article

# Retrospective Illumination Correction of Retinal Images

**Libor Kubecka, Jiri Jan, and Radim Kolar**

*The Faculty of Electrical Engineering and Communication, Brno University of Technology, 60200 Brno, Czech Republic*

Correspondence should be addressed to Radim Kolar, kolarr@feec.vutbr.cz

Received 11 November 2009; Accepted 22 May 2010

Academic Editor: Jiang Hsieh

Copyright © 2010 Libor Kubecka et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A method for correction of nonhomogenous illumination based on optimization of parameters of B-spline shading model with respect to Shannon's entropy is presented. The evaluation of Shannon's entropy is based on Parzen windowing method (Mangin, 2000) with the spline-based shading model. This allows us to express the derivatives of the entropy criterion analytically, which enables efficient use of gradient-based optimization algorithms. Seven different gradient- and nongradient-based optimization algorithms were initially tested on a set of 40 simulated retinal images, generated by a model of the respective image acquisition system. Among the tested optimizers, the gradient-based optimizer with varying step has shown to have the fastest convergence while providing the best precision. The final algorithm proved to be able of suppressing approximately 70% of the artificially introduced non-homogenous illumination. To assess the practical utility of the method, it was qualitatively tested on a set of 336 real retinal images; it proved the ability of eliminating the illumination inhomogeneity substantially in most of cases. The application field of this method is especially in preprocessing of retinal images, as preparation for reliable segmentation or registration.

## 1. Introduction

Improper scene illumination as well as nonideal acquisition conditions due to for example, misadjusted imaging system can introduce severe distortions into the resulting image. These distortions are usually perceived as smooth intensity variations across the image. According to the terminology commonly used in processing of magnetic resonance (MR) images, we call these systematic intensity level inhomogeneities as the *bias field*. With such unevenness, the subsequent image processing like image registration, segmentation, or pattern recognition may be substantially complicated; therefore, the correction of illumination inhomogeneities is highly desirable. Unfortunately, separating the bias field from the true underlying image is an underconstrained and ill-posed problem with fewer observations than free variables. For this reason, regularization of the problem is necessary.

Most existing bias correction methods assume that the bias field is multiplicative, slowly varying, and tissue independent. Many techniques ignore the noise and apply a log transform to make the bias field additive. The known illumination correction methods can be categorized in

the following groups: filtering, segmentation based, surface fitting, and other methods.

Illumination inhomogeneities are generated during acquisition process in systems with different modalities. Here, the proposed illumination correction method will be applied on retinal images from confocal scanning laser ophthalmoscope (CSLO). This correction is an important preprocessing task in image segmentation and/or multimodal registration [1–4].

The earliest bias correction techniques were based on phantoms [5] with known structure; this approach enables estimation of the bias field, its inversion and removal. Provided the calibrated phantom is available, this method can be considered the ground truth but in real situations is not often applicable.

Linear filtering methods [6, 7] try to estimate and eliminate the additive bias of the image using unsharp masking with defocus larger than any object in the image. These techniques can be extended, via nonlinear morphological approach to the multiplicative bias field estimation [6]. Homomorphic unsharp filtering methods [8–10] assume a separation of the low frequency bias field from the higher frequencies of the image structures.

Some simple methods, as [11], require expert supervision, which is time-consuming and often too subjective. In this approach, the expert specifies some parameters related to the intensity models of the different tissues and forms a list of intensity values and locations related to the background or to object classes. Then, the bias field over the image may be obtained by least-squares fitting to the intensity values at the preselected points. These methods are closely connected to segmentation-based methods. Authors of [12] combine shading correction with adaptive segmentation. They use a fuzzy C-means algorithm to allow labeling of a pixel to be influenced by its immediate neighbors. Main benefits of this approach are high robustness to salt-and-pepper noise and computational efficiency of the algorithm.

The expectation maximization (EM) algorithm is proposed in [13] to compute iteratively the optimal smooth bias field corrupting the data based on classification into several tissue classes. Their formulation includes the bias distortion in the statistical model of the pixel distribution, that is, the bias field influences the distribution by locally changing its mean value. The algorithm iterates two steps, the E-step calculating the posterior tissue probabilities, and the M-step estimating the bias field. The Styner's method [14] is based on a simplified model of the imaging process, a parametric model of a tissue class statistics, and a polynomial model of the inhomogeneity field. The estimation of the parametric bias field is formulated as a non-linear energy minimization problem solved by an evolution strategy.

Segmentation-based approaches raise the problem of selecting the number of classes, which have to be explicitly modeled. Furthermore, these algorithms unfortunately tend to converge to a local nonoptimal minimum for some bias configurations, especially when more than two tissue classes are modeled [15]. The main problem of segmentation-based technique is that the determination of specific class-conditioned intensity is difficult when the image is incorrectly illuminated; however, the illumination correction is the main aim. Also, these methods usually assume that the intensity distribution of an image is normal and given by distributions of individual tissues, which may be often invalid. The above drawbacks are even more serious when correcting pathological image data. Therefore, Likar et al. [16] used the closed connection between intensity nonuniformity correction and segmentation, and proposed a method, which iteratively interleaves them, so that both of them gradually improve, until the final correction and segmentation is reached. The method is based on iterative minimization of the class square-error of intensity distributions caused by non-uniformity and on a nonparametric segmentation method. The method makes no assumption on the distribution of tissue intensity and does not require initialization.

To overcome the segmentation problems, bias correction methods not requiring the segmentation were designed based on a chosen image quality criterion. In [17], a presumed histogram matching method is presented. Authors of [18] suggested the iterative optimization method, which seeks the smooth multiplicative field that maximizes the higher frequency content of the distribution of tissue

intensity. It requires a parametric model for the bias field but not a decomposition of the intensities. The quality criterion derived from the information theory has been also used in several applications. In [19], a polynomial multiplicative and additive shading model was introduced employing the cost-function based on image entropy and Powell optimizer. Mangin [20] uses image entropy combined with a measure of the field smoothness as the image quality cost function. Authors of [14] model the bias field using Legendre polynomials and use the energy function based on a multiclass model estimator and evolutionary search algorithm.

Specific methods of illumination correction were proposed in the frame of retinal image processing and analysis. Simple and fast methods using large-kernel median filter to obtain a low-pass correction coefficients were used for CSLO image preprocessing in [21]. Authors of [22] model the bias field (background image) of a fundus (basic retinal) image as a white Gaussian random field and use Mahalanobis distance for background pixel classification. Contrast normalization using high-pass filtered image is used in [2] as one step of microaneurysm detection procedure. Additive model of nonuniform illumination is used in [3], together with adaptive histogram equalization.

Other approaches exist, for example, in applications including illumination correction for face recognition [23, 24] and restoration of digitized documents [25, 26]. The latter application frequently uses multiplicative model and specific properties of the processed images (e.g., illumination edges or geometrical distortion).

In this paper, we focus our attention on methods estimating the parametric illumination field using the quality criterion derived from the information theory. We use a multiplicative model of nonuniform illumination and parametric local bias model for formulation of criterion function and its derivatives (Section 2). The results are presented in Section 3 for different optimizers, which were tested for two image sets—artificially illuminated images and real CSLO images. The assessment of the results concludes the paper in Section 4.

## 2. Methods

**2.1. Acquisition Model.** For the purpose of illumination correction process, we need a model of image creation. We assume, that each tissue class (vessels, optic disc, retinal surface) has a different mean value  $\rho(\mathbf{x})$  of the property measured by the imaging device. Moreover, every class of tissue has a characteristic texture, which can be modeled by the additive noise  $n_{\text{tiss}}(\mathbf{x})$ . The ideal output signal  $o(\mathbf{x})$  therefore consists of piecewise constant values plus additive noise. Due to finite size of the point spread function  $h(\mathbf{x})$  of the imaging device (different from the ideal Dirac impulse), this ideal signal is corrupted by convolution with  $h(\mathbf{x})$  and with the noise generated by the device  $n(\mathbf{x})$  (thermal or electronic noise and the noise coming from digitization during acquisition process). The overall equation of the observed image data can be formalized as follows ( $b(\mathbf{x})$  being

the illumination), see Figure 1:

$$\begin{aligned} s(\mathbf{x}) &= [(\rho(\mathbf{x}) + n_{\text{tiss}}(\mathbf{x}))b(\mathbf{x})]^* h(\mathbf{x}) \\ + n(\mathbf{x}) &\approx o(\mathbf{x})b(\mathbf{x}) + n(\mathbf{x}). \end{aligned} \quad (1)$$

On the right-hand side of (1), we neglected the smoothing due to the imaging properties not playing a substantial role in our problem (the disturbance need not be taken into account in the retinal applications).

Thus, under the assumption that the bias involved in image creation process is multiplicative, and that it is the only important disturbance, we can reconstruct the original signal as

$$\hat{o}(\mathbf{x}) = \frac{s(\mathbf{x})}{\hat{b}(\mathbf{x} | \Phi)} - \frac{n(\mathbf{x})}{\hat{b}(\mathbf{x} | \Phi)} \approx \frac{s(\mathbf{x})}{\hat{b}(\mathbf{x} | \Phi)} = s(\mathbf{x})\hat{b}^{-1}(\mathbf{x} | \Phi), \quad (2)$$

where  $\hat{b}(\mathbf{x})$  is the optimal illumination bias model controlled by parameters  $\Phi_i$  forming the vector  $\Phi$ , while  $\hat{b}^{-1}(\mathbf{x})$  is its reciprocal value. Example of a line input signal (measured line profile) can be seen on Figure 2. The choice of a proper bias model  $b(\mathbf{x})$  as well as of an appropriate criterion function, evaluating how well the non-homogenous illumination of the image is corrected for current values of the bias model parameters, is crucial.

**2.2. Bias Model.** In this section, we extend the Likar's algorithm [19] and applied the modified algorithm for retrospective shading correction of retinal images. We derive an expression for the criterion of quality of illumination compensation based on Shannon's entropy and on Parzen windowing probability estimation. Further, as a novelty, we derive analytical expressions for derivatives of this criterion with respect to parameters of the used B-spline multiplicative illumination model. The analytical expressions ease substantially the optimization calculations compared to purely numerical evaluation of derivatives.

Generally, in accordance with [14, 19], the intensity transformation performed by the reciprocal bias model can be defined by a linear combination of  $K$  smooth basis functions  $r_i(\mathbf{x})$ .

$$\hat{b}^{-1}(\mathbf{x}) = \sum_{i=1}^K \Phi_i r_i(\mathbf{x}), \quad (3)$$

where  $\Phi_i$  are the parameters of the transform defining the contribution of each basis.

In order to regularize the problem of finding the bias field that would optimally correct the illumination inhomogeneity, we used the bases in the following form:

$$r_i(\mathbf{x}) = \frac{q_i(\mathbf{x}) - {}^m c(\mathbf{x})}{{}^n c(\mathbf{x})}, \quad (4)$$

where  $q_i(\mathbf{x})$  are smooth polynomial basis functions,  ${}^m c(\mathbf{x})$  is a mean correction coefficient, and  ${}^n c(\mathbf{x})$  is a normalization coefficient. The correction coefficients are defined using constraints presented in [19], where the mean preserving

condition preventing a global shift of the mean intensity of the resultant corrected image is formulated as

$$\frac{1}{\Theta} \sum_{\mathbf{x} \in \Omega} s(\mathbf{x}) = \frac{1}{\Theta} \sum_{\mathbf{x} \in \Omega} s(\mathbf{x}) \hat{b}^{-1}(\mathbf{x}). \quad (5)$$

Here,  $\Theta$  is the size of the image domain  $\Omega$  (or region of interest: ROI) in pixels. Further, as we use the multiplicative bias model, the final intensity is differently sensitive to values of different parameters, thus producing non-linear logarithmic scale (see [19] for detailed description). This fact is undesirable in optimization. Therefore, to assure that every parameter has the same influence on the intensity transform, a normalization constraint for every base function kernel  $r_i(\mathbf{x})$  is introduced as

$$\frac{1}{\Theta} \sum_{\mathbf{x} \in \Omega} |s(\mathbf{x}) r_i(\mathbf{x})| = 1, \quad \forall i, \quad (6)$$

We have found that the illumination distortion of retinal images shows a significant spatial variance, while both illumination models mentioned in [14, 19] have rather a global character influencing the whole image despite possibly only locally defined distortions (e.g., when the image is corrupted in its upper left corner only, the polynomial model may significantly influence the opposite corner as well). Therefore, to model such local distortions, high-order models have to be used, which results in a high number of parameters to optimize, besides a danger of oscillatory character of the functions.

Hence, we decided to use the locally defined mean-corrected and normalized reciprocal bias model based on B-splines, formalized as follows:

$$\begin{aligned} \hat{b}^{-1}(\mathbf{x}) &= 1 + \sum_{\mathbf{x}_i \in P} \Phi(\mathbf{x}_i) \left[ \frac{(\beta^{(d)}(\Delta \mathbf{x}_i) - {}^m c(\mathbf{x}_i))}{{}^n c(\mathbf{x}_i)} \right], \\ \text{where } \Delta \mathbf{x}_i &= \frac{\mathbf{x} - \mathbf{x}_i}{\mathbf{h}}. \end{aligned} \quad (7)$$

For the case of two dimensions,  $\beta^{(d)}(\mathbf{x}) = \beta^{(d)}(x_1) \cdot \beta^{(d)}(x_2)$  is a separable  $d$ th order B-spline kernel,  $P$  is a set of  $n_{x1} \times n_{x2}$  control points of the transform regularly deployed over the extended image domain  $\Omega$  with the spacing  $\mathbf{h} = [h_1, h_2]$  ( $h_k$  possibly different but constant for each dimension, see Figure 3 for illustrative description). The division indicated in the expression for  $\Delta \mathbf{x}$  in (7) should be understood element-wise; thus  $\mathbf{x}_i/\mathbf{h}$  are the integer control-point indices. The  $\Phi(\mathbf{x}_i) = \phi_i$  are scalar coefficients corresponding to the control point. These coefficients are the parameters of the bias model and express the weight of influence of each B-spline basis function. As we use the 3rd order B-spline, which is nonzero only for  $\Delta x_k < -2, 2 > \epsilon$ , only  $4 \times 4$  basis functions influence the computation of a current point  $\mathbf{x}$  for two dimensions. For efficient implementation, the grid of control points has to be extended beyond the image borders.  ${}^m c(\mathbf{x}_i) = {}^m c_i$  are coefficients ensuring the mean conservation condition (5) and  ${}^n c(\mathbf{x}_i) = {}^n c_i$  are coefficients ensuring the normalization of the parameters (6); both these coefficients are precomputed ahead of the shading

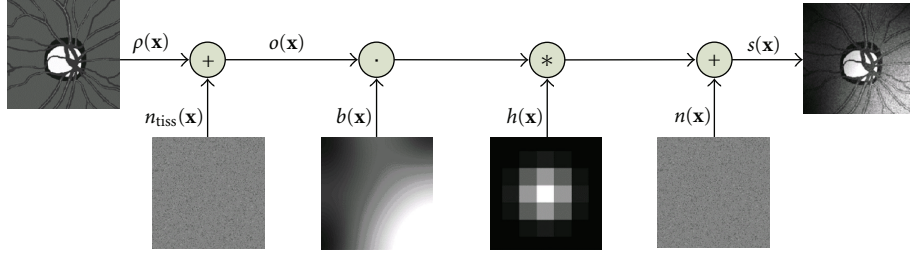


FIGURE 1: Model of image acquisition. Ideal image is corrupted by noise, multiplicative illumination, and finite resolution of the imaging device (characterized by Gaussian point spread function).

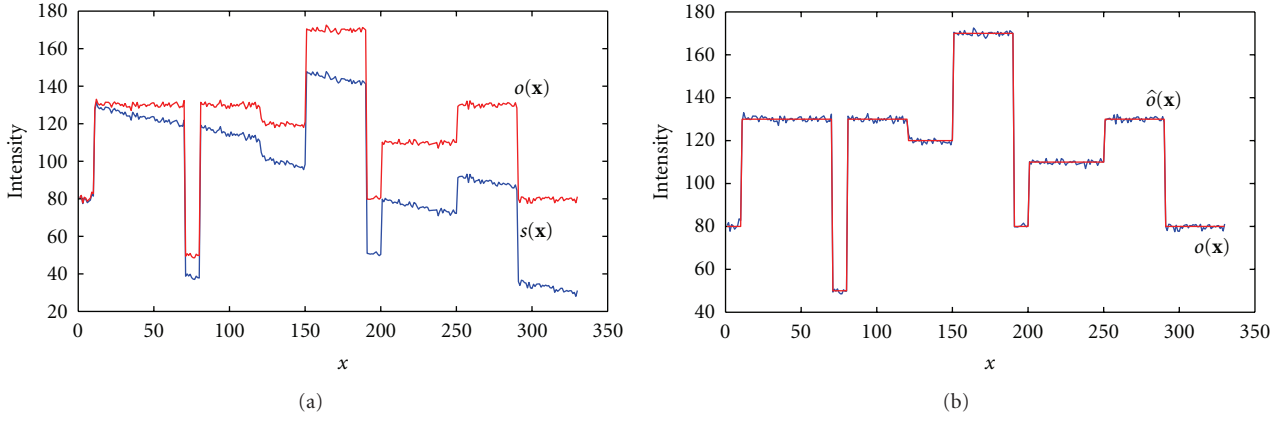


FIGURE 2: (a): Example of a line of input signal:  $o(\mathbf{x})$  is the ideal signal with the texture noise,  $s(\mathbf{x})$  is acquired signal furthermore corrupted by non-homogenous illumination (b): results of the correction process,  $o(\mathbf{x})$ —ideal signal,  $\hat{o}(\mathbf{x})$ —signal after entropy-based correction (nearly ideally recovered).

correction. From the mean preserving condition (5) and from (7), we have

$$\frac{1}{\Theta} \sum_{\mathbf{x} \in \Omega} s(\mathbf{x}) = \frac{1}{\Theta} \sum_{\mathbf{x} \in \Omega} \left\{ s(\mathbf{x}) \left[ 1 + \sum_{i=1}^K \frac{\Phi_i (\beta^{(d)}(\Delta \mathbf{x}_i) - m_{c_i})}{n_{c_i}} \right] \right\}, \quad (8)$$

which can be configured into

$$\frac{1}{\Theta} \sum_{i=1}^K \left\{ \frac{\Phi_i}{n_{c_i}} \left[ \sum_{\mathbf{x} \in \Omega} s(\mathbf{x}) (\beta^{(d)}(\Delta \mathbf{x}_i) - m_{c_i}) \right] \right\} = 0. \quad (9)$$

The nontrivial solution ( $\Phi_i \neq 0$ ) of (9) is

$$\sum_{\mathbf{x} \in \Omega} s(\mathbf{x}) (\beta^{(d)}(\Delta \mathbf{x}_i) - m_{c_i}) = 0, \quad \forall i, \quad (10)$$

which provides

$$m_{c_i} = \frac{\sum_{\mathbf{x} \in \Omega} s(\mathbf{x}) \beta^{(d)}(\Delta \mathbf{x}_i)}{\sum_{\mathbf{x} \in \Omega} s(\mathbf{x})} \quad (11)$$

for each mean preserving constant  $m_{c_i}$  corresponding to a control point  $i$  with coordinates  $\mathbf{x}_i$ .

Similarly, we can derive the normalization condition from (6) and (7)

$$\frac{1}{\Theta} \sum_{\mathbf{x} \in \Omega} \left| \frac{s(\mathbf{x}) (\beta^{(d)}(\Delta \mathbf{x}_i) - m_{c_i})}{n_{c_i}} \right| = 1, \quad \forall i \quad (12)$$

from where

$$n_{c_i} = \frac{1}{\Theta} \sum_{\mathbf{x} \in \Omega} |s(\mathbf{x}) (\beta^{(d)}(\Delta \mathbf{x}_i) - m_{c_i})|, \quad (13)$$

for each normalization constant  $n_{c_i}$ . Details of normalization coefficients computation in case of polynomial basis can be found in [19]; here we modified this approach for the B-spline representation. Finally, the restored image  $\hat{o}(\mathbf{x})$  is provided using (2) by the intensity transform  $\hat{b}^{-1}(\mathbf{x} | \Phi)$  with the found parameter vector  $\Phi$ .

Because this transformation may generally produce out-of-range intensity values, we use intermediate image representation with extended intensity range. The resultant image is finally computed using linear contrast transformation converting the intermediate image back into the original intensity range.

**2.3. Criterion Function.** In order to find parameters of the bias model describing the undesirable illumination, we need to define a criterion, with respect to which the parameters would be optimized. In the works of Likar et al. [19] and Mangin [20], the image entropy is shown to be a suitable criterion. Their idea is that the illumination is additional information added to the information included in the original signal  $o(\mathbf{x})$  and because we would like to remove the illumination bias, the information content of the corrected

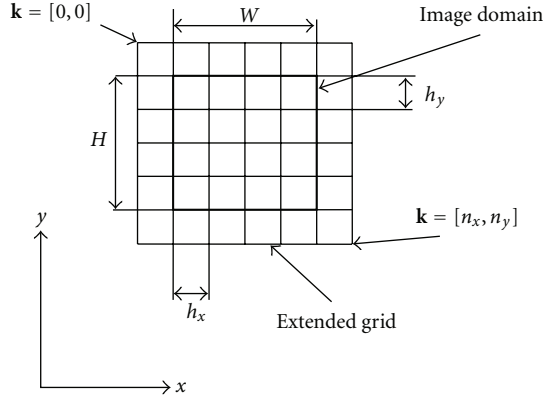


FIGURE 3: Formation of grid of control points for B-spline intensity transform.

image should be lower than that of the distorted image  $s(\mathbf{x})$ . Therefore, when looking for parameters of the reciprocal bias model eliminating the non-homogenous illumination, the Shannon's entropy  $H(\cdot)$  of the resulting image  $\hat{o}(\mathbf{x})$ ,

$$H = -\sum_k P(k) \log(P(k)), \quad (14)$$

which is a measure of the information amount, should be minimized. Here,  $P(k)$  is the probability of intensity  $k$  appearing in any pixel of  $\hat{o}(\mathbf{x})$ .

Although the brightness  $k$  is a discrete variable in reality (represented by 8 bits, i.e., 256 values), it may be well approximated by continuous variable  $\kappa$  with a probability density  $p(\kappa)$ . Then, in order to analytically derive the derivatives of the criterion, we may assume that the amount of information can be described by the integral version of (14) for the continuous variable  $\kappa$  as

$$H = -\int_{\kappa} p(\kappa) \log(p(\kappa)) d\kappa. \quad (15)$$

However, the probability density  $p(\kappa)$  is not available explicitly and must be estimated from the image data transformed by the current parameters  $\Phi$ , providing that the image has been generated by a homogeneous stochastic field. For this purpose, we used the Parzen windows (PW) technique [20], also known as kernel density estimator. In this scheme, the density is constructed by taking intensity samples  $\hat{o}(\mathbf{x}; \Phi)$  from the transformed image and super-positioning kernel functions  $\beta^{(3)}$  centered on the elements of  $\Omega$  as illustrated in the Figure 4. More formally,

$$p(\kappa; \Phi) = \frac{1}{\Theta} \sum_{\mathbf{x} \in \Omega} \beta^{(3)}\left(\frac{\hat{o}(\mathbf{x}; \Phi) - \kappa}{z}\right) \quad (16)$$

is the probability density estimate at intensity value  $\kappa$ , which is obtained using PW,  $\beta^{(3)}$  is 3rd order B-spline kernel,  $z$  is a constant parameter defining the width of the intensity class by controlling width of the super-positioned B-spline kernel (analogue to histogram bin size), and  $\Theta$  is a normalization coefficient assuring that the integral of  $p(\kappa)$  is equal to

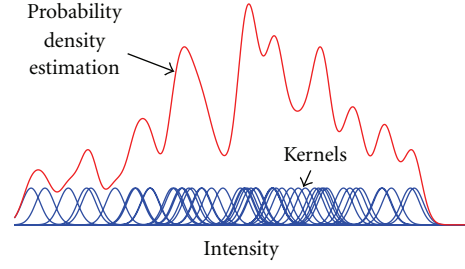


FIGURE 4: Illustration of Parzen windowing method for estimating density probability of image intensities.

one. This method can be looked at like at a convolution of impulses at different intensities with a spline kernel (see Figure 4). The advantage of this method is the possibility of expressing the probability derivatives analytically and also the possibility to speed-up the computation by taking into account only a subset of the whole sample set defined in the region of interest  $\Omega$ .

Thanks to our formulation of the criterion (15) based on PW, both the probability estimation and the intensity transformation are defined continuously and therefore we can analytically derive partial derivatives of the criterion with respect to components of the parameter vector  $\Phi$ . As  $H$  is a compound function of the form

$$H(\kappa; \Phi) = f(p(\beta(\hat{o}(\Phi)))), \quad (17)$$

its derivative with respect to an element of  $\Phi$  can be expressed as

$$\frac{\partial H(\kappa; \Phi)}{\partial \Phi_k} = \frac{\partial H}{\partial p} \frac{\partial p}{\partial \beta} \frac{\partial \beta}{\partial \hat{o}} \frac{\partial \hat{o}}{\partial \Phi_k}. \quad (18)$$

From here on, the  $i$ th control point (a linearly decoded control point of the rectangular grid) is characterized by a vector index  $\mathbf{k} = [k_1, k_2]$  with appropriate indices along both axes. If  $n_{x1}$  is the number of control points along the first dimension, then  $i = k_1 + n_{x1}k_2$ .

By applying product rule, we obtain from (15)

$$\frac{\partial H(\kappa; \Phi)}{\partial p} = -\int_{\kappa} (\log(p(\kappa; \Phi)) + 1) d\kappa. \quad (19)$$

Calculation of this integral is approximated using the rectangle rule as

$$\frac{\partial H(\kappa; \Phi)}{\partial p} \approx -\sum_k (\log(p(\kappa_k; \Phi)) + 1). \quad (20)$$

Further from (16),

$$\frac{\partial p}{\partial \beta} = \sum_{\mathbf{x} \in \Omega} \frac{1}{\Theta}. \quad (21)$$

The derivative of the B-spline kernel can be expressed using B-spline properties [27] as

$$\frac{\partial \beta^{(3)}(\xi)}{\partial \xi} = \beta^{(2)}\left(\xi + \frac{1}{2}\right) - \beta^{(2)}\left(\xi - \frac{1}{2}\right). \quad (22)$$



The image  $\hat{o}(\mathbf{x}; \Phi)$  derived from the observed image  $s(\mathbf{x})$  by application of the intensity transform with parameters  $\Phi$  can be expressed using (2) and (7) as

$$\hat{o}(\mathbf{x}; \Phi) = s(\mathbf{x}) \left\{ 1 + \sum_{\mathbf{x}_i \in P} \Phi(\mathbf{x}_i) \left[ \frac{\beta^{(n)}(\Delta \mathbf{x}_i) - {}^m c(\mathbf{x}_i)}{{}^m c(\mathbf{x}_i)} \right] \right\},$$

where  $\Delta \mathbf{x}_i = \frac{\mathbf{x} - \mathbf{x}_i}{\mathbf{h}}$

(23)

and for its partial derivatives we can write (in two dimensions):

$$\frac{\partial \hat{o}(\mathbf{x}; \Phi)}{\partial \Phi_k} = \sum_{\mathbf{x}_i \in P} \left\{ s(\mathbf{x}) \frac{1}{{}^n c_i} \left[ \beta^{(3)} \left( \frac{x_1 - x_{k_1}}{h_1} \right) - {}^m c_i \right] \right. \\ \left. \times \frac{1}{{}^n c_i} \left[ \beta^{(3)} \left( \frac{x_2 - x_{k_2}}{h_2} \right) - {}^m c_i \right] \right\}.$$
(24)

Finally, we have for the derivatives of the criterion

$$\frac{\partial H(\kappa; \Phi)}{\partial \Phi_k} \approx - \sum_{\kappa} \frac{\partial p(\kappa; \Phi)}{\partial \Phi_k} (\log(p(\kappa; \Phi)) + 1),$$
(25)

where

$$\frac{\partial p(\kappa; \Phi)}{\partial \Phi_k} = \sum_{\mathbf{x} \in \Omega} \frac{1}{\Theta} \left\{ \sum_{\mathbf{x}_i \in P} \frac{\partial \beta^{(3)}(\xi)}{\partial \xi} \right\}_{\xi = (s(\mathbf{x})b^{-1}(\mathbf{x})/z) - \kappa} \\ \times s(\mathbf{x}) \frac{1}{{}^n c_i} \left[ \beta^{(3)} \left( \frac{x_1 - x_{k_1}}{h_1} \right) - {}^m c_i \right] \frac{1}{{}^n c_i} \\ \times \left[ \beta^{(3)} \left( \frac{x_2 - x_{k_2}}{h_2} \right) - {}^m c_i \right] \}.$$
(26)

The grid controlling the compensation field is chosen adequately sparse as required by the smooth character of the to be compensated illumination unevenness. An example illustrating behaviour of the criterion function  $H(\kappa, \Phi)$  as dependent on one element of the parameter vector—concretely  $\Phi_{[0,2]}$  at  $i = 2$ , that is, for  $k = [0,2]$ —top-right image corner on  $3 \times 3$  grid covering the image—is depicted on Figure 5(a) showing a distinct minimum. Below, also the course of partial derivative of  $H(\kappa, \Phi)$  is shown, both as derived analytically by the described approach and as calculated numerically via finite differences. There is a good agreement between both versions namely, in the important area around the; however, important advantages of analytical derivatives are faster evaluation and smooth behavior in the parameter space. On Figure 5(b), the influence of this parameter on the illumination correction is illustrated for five particular values of  $\Phi_{[0,2]}$  influencing namely the correction in the upper right corner of the image; obviously, the best corrected image corresponds to the minimum of the criterion thus to the zero position of its derivative.

**2.4. Optimization.** The overall illumination correction algorithm is based on the reciprocal illumination model  $b^{-1}$  with the parameters minimizing the entropy criterion  $H$ . The optimization aims at finding the optimum parameter vector

$$\Phi_{\text{opt}} = \arg \min_{\Phi} \{H(s(\mathbf{x})b^{-1}(\mathbf{x}))\}. \quad (27)$$

The block diagram of the iterative algorithm is depicted on Figure 6.

Various types of optimization techniques were studied in order to find the optimizer best suited for the particular properties of the used optimization criterion in the problem of non-homogenous illumination correction. The tested methods were namely, downhill simplex [28] (Amoeba), Powell's direction set [28], controlled random search (CRS) [29], gradient descent (GD) [28, 30], conjugate gradient (CG) [28], and two versions of the limited memory Broyden, Fletcher, Goldfarb, Shannon (LBFGS) methods [31].

The comparison of results of the individual methods can be found in the next paragraph.

### 3. Experiments and Results

The proposed algorithm is supposed to be able to deal with non-homogenous illumination of retinal image data ( $384 \times 384$  pixel,  $10 \mu\text{m}/\text{pixel}$ ) obtained by means of the CSLO (Heidelberg Retinal Tomograph HRT II). In order to check the efficiency of the designed method on images with a known illumination inhomogeneity, a model of the HRT II image signal was designed, using segmentation of real HRT II images into five object classes with the estimated characteristic intensity values assigned to the objects as follows.

0—black background introduced by HRT II; 60—vessels; 120—vessel centers (brighter due to reflections); 90—retinal tissue; 30—optic disc; 250—optic disc cup). Further, according to the image acquisition model (1), random tissue noise  $n_{\text{tiss}}$  (uniformly distributed with standard deviation  $\sigma = 28$ ) was added and the resulting image was convolved with the estimated point spread function of the real imaging system, approximated by 2D Gaussian kernel with standard deviation  $\sigma = 1$ . Finally, an artificial multiplicative illumination field controlled by a mesh of  $3 \times 3$  nodes on the image area with random parameters was applied (see Figure 7).

A set of 40 simulated images was created via this model. Normalized parameters of the illumination field were uniformly distributed on interval  $[-15, 15]$ . Then, various optimizers were tested and evaluated with respect to achieved quality of the retrospective compensation of the illumination unevenness. The compensation quality was quantified using the fact that the known artificially introduced multiplicative illumination field and the final illumination correction field should be ideally inverse. Therefore, we can define the postcorrection illumination error  $\xi_{\text{post}}$  as

$$\xi_{\text{post}} = \frac{1}{\Theta} \sum_{\mathbf{x} \in \Omega} \sqrt{(1 - b(\mathbf{x})\hat{b}^{-1}(\mathbf{x}))^2}, \quad (28)$$

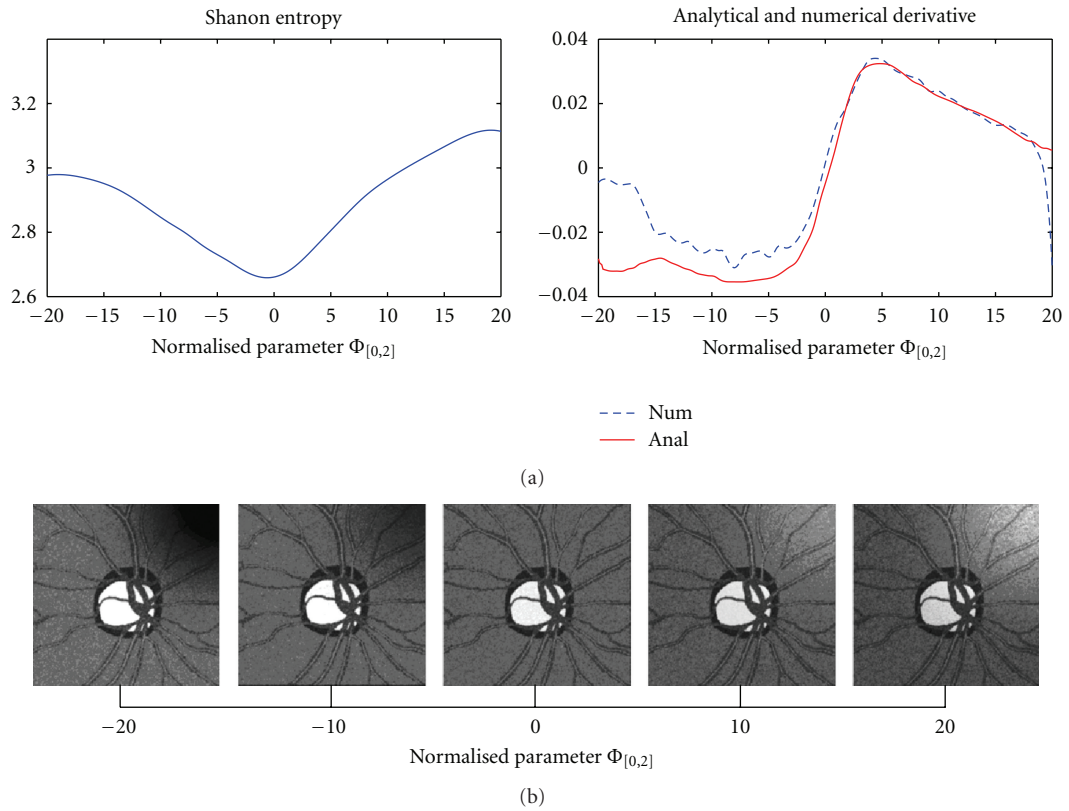


FIGURE 5: (a): Shannon entropy-based criterion and its partial derivative as functions of the parameter  $\Phi_{[0,2]}$  (control point in upper-left image corner) (b): Examples of influence of  $\Phi_{[0,2]}$  varying from -20 to 20 with step 10, which primarily influences the upper right corner correction; notice the optimum illumination correction for the value of  $\Phi_{[0,2]}$  corresponding to the zero derivative position

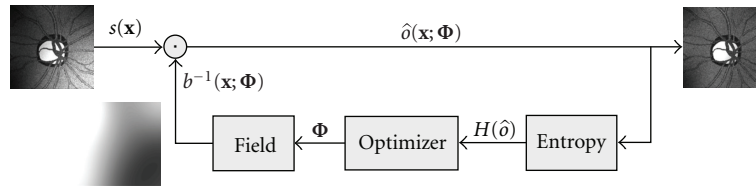


FIGURE 6: Flow diagram of iterative optimization algorithm.

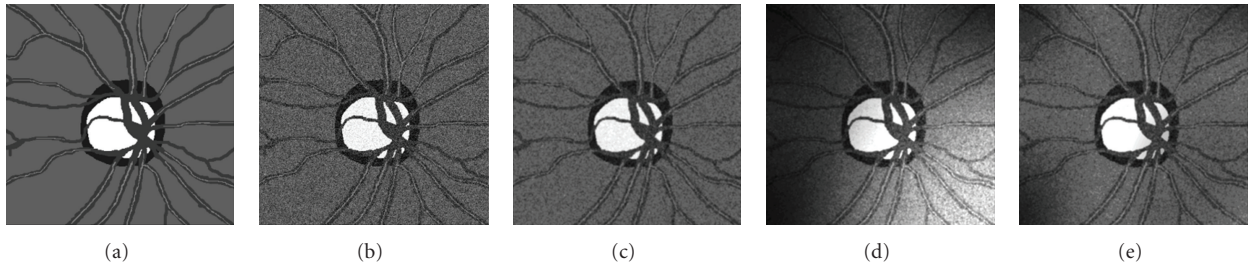


FIGURE 7: (a): Ideal model image created by manual segmentation of a real image, (b): Model image after noise addition, (c): Model image after noise addition and convolution with PSF of the imaging system. (d): Final model image corrupted by non-homogenous illumination. (e): the image after processing by the presented algorithm.

TABLE 1: Study of suitability of different optimizers for the task of optimization of Shannon's entropy. The best values are highlighted.

	Nongradient-based methods			Gradient-based methods			
	Amoeba	Powell	CRS	L-BFGS	L-BFGS-B	<b>GD</b>	CG
$\xi_{post}$	0.16	0.23	0.14	0.13	0.14	<b>0.09</b>	0.18
$t$ [s]	68.6	95.6	514	14.7	<b>15.5</b>	45.9	410
N	163	226	1201	26.4	<b>21.0</b>	57.4	615
$\rho$ [%]	44.5	18.6	52.9	55.9	51.3	<b>67.7</b>	36.9

$\xi_{pre} = 0.28$ .

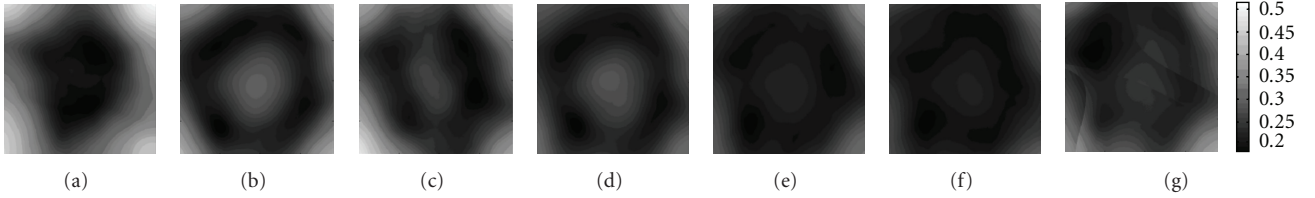


FIGURE 8: (a) Artificially introduced illumination field. Error illumination field (illumination and restoration field multiplied, ideally black) using: (b) Amoeba optimizer, (c) Powel optimizer, (d) CRS, (e) LBFGS optimizer, (f) Gradient descent optimizer (g) Conjugate gradient optimizer.

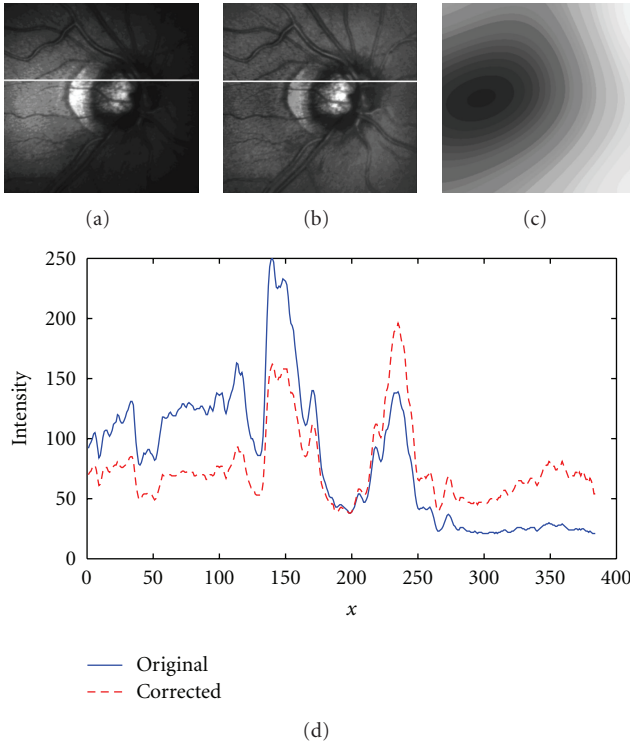


FIGURE 9: (a) Retinal image highly corrupted by non-homogenous illumination, (b) Image after multiplicative correction by recovered bias field, (c) Normalized bias field controlled by  $3 \times 3$  parameters automatically obtained using proposed algorithm, (d) Intensity profiles along the indicated row.

where  $b(\mathbf{x})$  is the introduced illumination field,  $\hat{b}^{-1}(\mathbf{x})$  is the found correction of the illumination, and  $\Theta$  is the size of the image domain  $\Omega$ . The mean precorrection illumination error

TABLE 2: Influence of different number of samples used for entropy evaluation on the algorithm precision and speed.

	GD, 50 bins		
coverage	0.15	0.5	1
$\xi_{post}$	<b>0.09</b>	<b>0.09</b>	<b>0.09</b>
$t$ [s]	45.8	46.1	212
$N$	57.4	57.5	67.6

defined as

$$\xi_{pre} = \frac{1}{\Theta} \sum_{\mathbf{x} \in \Omega} \sqrt{(1 - b(\mathbf{x}))^2} \quad (29)$$

was set to  $\xi_{pre} = 0.2875$ , that is, the intensity of the ideal image was corrupted on average by approximately 30%.

The seven optimization algorithms mentioned in the previous section were tested and the average results obtained when using the simulated image set are summarized in Table 1. Here, the achieved after-correction total mean errors are presented, together with the computing times per image, numbers of iterations, and the achieved percentage suppression of the illumination unevenness. Examples of error fields resulting from these optimizers can be seen on Figure 8. The Amoeba, Powell and CRS methods need to evaluate the criterion value only; the derivatives of the criterion need not be computed. On the other hand, these three methods converge slowly and the total computational demands are higher compared to GD, CG, and LBFGS. Moreover, the worst problem of Amoeba, CRS, and especially of Powell algorithm was the unreliable convergence. The Powell optimizer was usually unable to find the optimum parameters of the correction illumination field (see Table 1). The false convergence to a side extreme is less frequent in case of smaller distortions introduced but even in this case the gradient-based algorithms LBFGS and GD outperform



the nongradient algorithms. This can be explained by smoother description of the shape of the multidimensional cost function when using the analytical derivatives. Surprisingly, the simplest gradient descent optimizer with variable step proved to be the best with respect to the correction precision, suppressing the illumination errors by about 67 per cent on average (from  $\xi_{\text{pre}} = 0.2875$  to  $\xi_{\text{post}} = 0.0926$ ). The LBFGS optimizer was the fastest but the average achieved suppression of the simulated illumination error was only 51 per cent.

In frame of experiments, we tested also the influence of different number of image samples used for probability approximation. As can be seen in Table 2, even as little as 15% (about 20 000) of the available samples are statistically powerful enough to provide good approximation of probability density of image intensities. This results in faster computation of the criterion.

The next set of experiments concerned real retinal images with clearly visible illumination inhomogeneities that however were not known. The algorithm was tested on the set of 336 real (clinically obtained) retinal images. Obviously, the quality evaluation of non-homogenous illumination compensation could only be performed subjectively, due to lack of the golden standard (i.e., ideal illumination in each case). In majority of cases (about 95%), a substantial improvement was visible; in the remaining cases, the image remained practically unchanged, that is, no image was further distorted by the correction. On Figure 9, a result of the designed algorithm as applied to a real retinal HRT II image is illustrated; the intensity profiles clearly show the successful bias correction.

## 4. Conclusions

A method for efficient illumination correction was proposed, implemented, and verified: quantitatively on simulated images with known deterioration and qualitatively on an extensive set of real retinal images lacking this knowledge. It is based on estimating a B-spline polynomial shading model, the inversion of which provides correction of the input image, optimal in sense of the used information-based criterion—Shannon's entropy. Previously published methods using a similar principle were modified namely, as to the type of the used correction function concerns, and as a novelty, the computation of the criterion is based on probability distribution estimates by Parzen windowing method, which enables us to derive analytical expressions for derivatives of the criterion. This consequently substantially speeds up the computations. Along with the efficient B-spline distortion model, it results in the possibility of efficient usage of gradient-based optimizers. We compared the efficiency and precision of three non-gradient and four gradient-based optimizers and found the classical gradient descent optimizer with variable step as the best for the purpose of illumination correction, formulated in the suggested way.

The quantitative tests were done on an image set artificially created with respect to characteristics of the

imaging device (the method is primarily aimed at improving quality of retinal images taken by means of the HRT II CSLO); these tests have shown that the designed algorithm is capable of removing up to about 70% of artificially introduced illumination variability. Finally, the method was successfully qualitatively tested on a set containing 336 real CSLO clinically obtained retinal images.

According to results of some further tests, involving subsequent registration of multiple images, preprocessing of the retinal images by the proposed algorithm had a clearly positive influence on reliability of the registration of the images using the registration method developed by our group [1], when compared to registration results concerning the unprocessed images. A similar positive effect can be expected for even higher types of image analysis following the presented correction method.

## Acknowledgments

This research has been supported by the National Research Centre DAR (sponsored by the Ministry of Education, Czech Republic) project no. 1M0572 and partly also by the research frame no. MSM 0021630513.

## References

- [1] L. Kubecka and J. Jan, "Registration of bimodal retinal images—improving modifications," in *Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '04)*, pp. 1695–1698, San Francisco, Calif, USA, September 2004.
- [2] A. D. Fleming, S. Philip, K. A. Goatman, J. A. Olson, and P. F. Sharp, "Automated microaneurysm detection using local contrast normalization and local vessel detection," *IEEE Transactions on Medical Imaging*, vol. 25, no. 9, pp. 1223–1232, 2006.
- [3] A. A.-H. Abdel-Razik Youssif, A. Z. Ghalwash, and A. A. S. Abdel-Rahman Ghoneim, "Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter," *IEEE Transactions on Medical Imaging*, vol. 27, no. 1, pp. 11–18, 2008.
- [4] R. Kolar, L. Kubecka, and J. Jan, "Registration and fusion of the autofluorescent and infrared retinal images," *International Journal of Biomedical Imaging*, vol. 2008, no. 1, Article ID 513478, 11 pages, 2008.
- [5] L. Axel, J. Costantini, and J. Listerud, "Intensity correction in surface-coil MR imaging," *American Journal of Roentgenology*, vol. 148, no. 2, pp. 418–420, 1987.
- [6] I. T. Young, "Shading correction: compensation for illumination and sensor inhomogeneities," in *Current Protocols in Cytometry*, J. P. Robinson, et al., Ed., vol. 1, pp. 2.11.1–2.11.12, John Wiley & Sons, New York, NY, USA, 2000.
- [7] R. Chrastek, G. Michelson, K. Donath, M. Wolf, and H. Niemann, "Vessel segmentation in retina scans," in *Proceedings of the 15th Biennial International EURASIP Conference Biosignal*, pp. 252–254, Brno, Czech Republic, 2000.
- [8] B. H. Brinkmann, A. Manduca, and R. A. Robb, "Optimized homomorphic unsharp masking for MR grayscale inhomogeneity correction," *IEEE Transactions on Medical Imaging*, vol. 17, no. 2, pp. 161–171, 1998.

- [9] R. B. Lufkin, T. Sharpless, B. Flannigan, and W. Hanafee, "Dynamic-range compression in surface-coil MRI," *American Journal of Roentgenology*, vol. 147, no. 2, pp. 379–382, 1986.
- [10] J. Haselgrove and M. Prammer, "An algorithm for compensation of surface-coil images for sensitivity of the surface coil," *Magnetic Resonance Imaging*, vol. 4, no. 6, pp. 469–472, 1986.
- [11] B. M. Dawant, A. P. Zijdenbos, and R. A. Margolin, "Correction of intensity variations in MR images for computer-aided tissue classification," *IEEE Transactions on Medical Imaging*, vol. 12, no. 4, pp. 770–781, 1993.
- [12] M. N. Ahmed, S. M. Yamany, A. A. Farag, and T. Moriarty, "Bias field estimation and adaptive segmentation of MRI data using a modified fuzzy c-means algorithm," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '99)*, vol. 1, pp. 250–255, Fort Collins, Colo, USA, June 1999.
- [13] W. M. Wells III, W. E. L. Crimson, R. Kikinis, and F. A. Jolesz, "Adaptive segmentation of mri data," *IEEE Transactions on Medical Imaging*, vol. 15, no. 4, pp. 429–442, 1996.
- [14] M. Styner, "Parametric estimate of intensity inhomogeneities applied to MRI," *IEEE Transactions on Medical Imaging*, vol. 19, no. 3, pp. 153–165, 2000.
- [15] R. Guillemaud and M. Brady, "Estimating the bias field of MR images," *IEEE Transactions on Medical Imaging*, vol. 16, no. 3, pp. 238–251, 1997.
- [16] B. Likar, J. Derganc, and F. Pernuš, "Segmentation-based retrospective correction of intensity non-uniformity in multi-spectral MR images," in *Medical Imaging 2002: Image Processing*, M. Sonka and J. M. Fitzpatrick, Eds., vol. 4684 of *Proceedings of SPIE*, pp. 1531–1540, San Diego, Calif, USA, February 2002.
- [17] L. Wang, H.-M. Lai, G. J. Barker, D. H. Miller, and P. S. Tofts, "Correction for variations in MRI scanner sensitivity in brain studies with histogram matching," *Magnetic Resonance in Medicine*, vol. 39, no. 2, pp. 322–327, 1998.
- [18] J. G. Sied, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in mri data," *IEEE Transactions on Medical Imaging*, vol. 17, no. 1, pp. 87–97, 1998.
- [19] B. Likar, J. B. A. Maintz, M. A. Viergever, and F. Pernuš, "Retrospective shading correction based on entropy minimization," *Journal of Microscopy*, vol. 197, no. 3, pp. 285–295, 2000.
- [20] J.-F. Mangin, "Entropy minimization for automatic correction of intensity nonuniformity," in *Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA '00)*, pp. 162–169, IEEE Press, Hilton Head Island, SC, USA, June 2000.
- [21] H. Niemann, R. Chrastek, B. Lausen et al., "Towards automated diagnostic evaluation of retina images," *Pattern Recognition and Image Analysis*, vol. 16, no. 4, pp. 671–676, 2006.
- [22] M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Medical Image Analysis*, vol. 9, no. 3, pp. 179–190, 2005.
- [23] T. Chen, W. Yin, X. S. Zhou, D. Comaniciu, and T. S. Huang, "Total variation models for variable lighting face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1519–1524, 2006.
- [24] J. Zhu, B. Liu, and S. C. Schwartz, "General illumination correction and its application to face normalization," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03)*, vol. 3, pp. 133–136, April 2003.
- [25] M. S. Brown, M. Sun, R. Yang, Y. Lin, and W. B. Seales, "Restoring 2D content from distorted documents," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 11, pp. 1904–1916, 2007.
- [26] G. Meng, N. Zheng, S. Du, Y. Song, and Y. Zhang, "Shading extraction and correction for scanned book images," *IEEE Signal Processing Letters*, vol. 15, pp. 849–852, 2008.
- [27] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing. Part I. Theory," *IEEE Transactions on Signal Processing*, vol. 41, no. 2, pp. 821–833, 1993.
- [28] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, Cambridge, UK, 1995.
- [29] M. M. Ali, A. Törn, and S. Viitanen, "A numerical comparison of some modified controlled random search algorithms," *Journal of Global Optimization*, vol. 11, no. 4, pp. 377–385, 1997.
- [30] D. P. Mandic, "A generalized normalized gradient descent algorithm," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 115–118, 2004.
- [31] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Mathematical Programming*, vol. 45, no. 1–3, pp. 503–528, 1989.