Journal of Multivariate Analysis 150 (2016) 191–213

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Illumination problems in digital images. A statistical point of view

S. Geffray [a,*], N. Klutchnikoff [b], M. Vimond [c]

[a] IRMA UMR 7501, Université de Strasbourg, 7 rue René Descartes, 67084 Strasbourg Cedex, France

[b] IRMAR CNRS UMR 6625, Université de Rennes 2, Rue du Recteur Paul Henry, 35000 Rennes, France

[c] CREST CNRS UMR 9194, Ensai, Campus de Ker-Lann, Rue Blaise Pascal - BP 37203, 35172 BRUZ cedex, France

ARTICLE INFO

Article history:
Received 20 March 2015
Available online 4 June 2016

AMS subject classifications:
62G08
62G20
62H35

Keywords:
Semiparametric estimation
Illumination problem
Denoising
Statistical image processing
Local polynomial regression

ABSTRACT

This paper focuses on the statistical treatment of illumination artefacts on digital images in the presence of an additional random noise. We assume that this artefact consists of "smooth" variations of the intensity of the signal of interest R. Such an assumption is classically modelled using a function L which acts in a multiplicative way on R. Our goal is to estimate R from observations of a random variable Y which obeys the regression model $Y = RL + \varepsilon$. Our main contribution lies in the derivation of a new estimator of R which is shown to be consistent under suitable identifiability and regularity conditions. The accuracy of this new estimation procedure is studied from a theoretical point of view through the rate of convergence of the uniform risk. Applications to real Scanning Electron Microscopy images are presented, as well as a qualitative study of the performances of our method with respect to other image processing techniques.

1. Introduction

In this paper, we are interested in a multi-dimensional signal which is observed in the presence of both an illumination artefact and an additional additive random noise. Illumination effects may arise in a variety of digital images including digital photography, digital microscopy such as Scanning Electron Microscopy (SEM) images (see Fig. 1) and Magnetic Resonance (MR) images. The illumination artefact considered here consists of colour/grey level intensity variations which are seen on the acquired image but which are not present in the original scene. This artefact is also termed as intensity nonuniformity, bias field, shading or gain field problem depending on the application domain. Such an artefact is obviously undesirable especially when automated measurements or quantitative analyses from digital images are the final goal; see e.g., [35] for immuno-histochemistry purposes. Such automated measurements from digital microscopy images are also frequently required in nano-polymer science; see for instance [2], [33] or [32]. Another instance of uneven illuminated image analysis comes from cell biology, see [64]. These authors develop an automated method of cell detection and counting from microscopic images. They specifically mention the fact that difficulties arise from an illumination gradient, the direction and intensity of which vary from one image to another due to experiment protocol and microscope operator variability. Whatever the context, illumination artefacts are inherent to the image acquisition techniques and cannot be removed by hardware changes during the acquisition process. In this paper, we aim at developing a consistent and reasonably simple procedure in order to get rid of both illumination artefacts and noise when an additional additive noise is present.

* Corresponding author.
E-mail address: geffray@math.unistra.fr (S. Geffray).

http://dx.doi.org/10.1016/j.jmva.2016.05.001

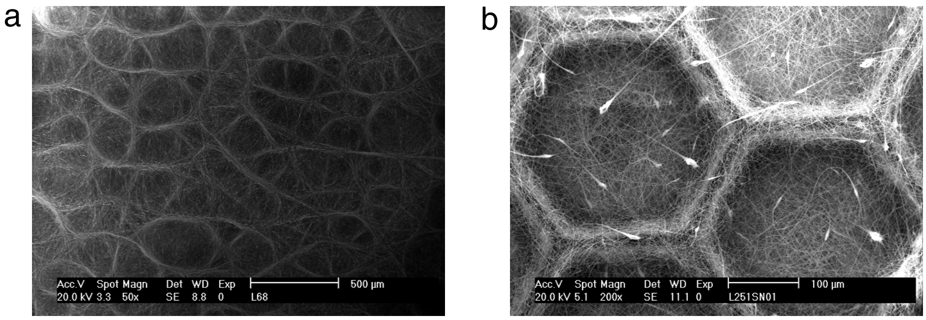

Fig. 1. Polymer fibre network images obtained by Scanning Electron Microscopy and suffering from illumination artefacts. Image courtesy by Guy Schlatter, ICPEES, Université de Strasbourg, France.

Illumination processing goes back to the pioneering work of Land and McCann [30]. Their so-called Retinex theory addresses both the colour constancy phenomenon and the human perception of illumination and reflectance, the latter being defined as the fraction of radiant light which is reflected from a surface. From there, an important amount of literature has dealt with the question of decomposing an image into illumination and reflectance. The most popular model consists of pixelwise multiplicative decomposition of any given image F into two different images, namely the reflectance image R, and the illumination image L. The problem is then to estimate/recover R from F. We refer the reader to the book by McCann and Rizzi [43] or to Agarwal et al. [1] for a complete exposition of the Retinex theory. Depending on the application at hand, different assumptions on the underlying reflectance image are to be fulfilled, and the resulting implementations vary significantly.

For instance, the most common MR image assumption is that the voxels corresponding to a given tissue type have similar intensities. Otherwise stated, the reflectance image is constant within finite regions of the image domain. This modelling approach is adopted e.g. in [50]. Within this framework, approaches taking into account tissue classes give very convincing results; see [4,19,65,26,31] and the references contained therein. The N4ITK method of Tustison et al. [63] is another proven track record method for bias correction in the noise-free case which is widely used in the MRI community. The bias field estimate is nonparametric in the sense that it does not rely on parametric segmentation hypotheses. The bias field is modelled with B-splines. The approach of the N4 method is then to sharpen iteratively the distribution of the observed image, and to use least-squares fitting to estimate the corresponding smooth field at each iteration.

So called natural images are of a rather different nature. In particular, the hypothesis that the reflectance is piecewise constant does not hold. Within the natural image framework, developments of Retinex methods have arisen until very recently; see e.g. [29,67] and reference therein. A statistical model appropriate to the reflectance–illumination problem is then a regression model where the reflectance is pointwise multiplicatively corrupted by the illumination with an additional additive noise. Fu et al. [15] consider such a model and propose a Bayesian inferential point of view. They set a prior on both the illumination image and its gradient. Within the framework of mixed models, Demidenko [8] estimates the light direction from a simple linear regression model and estimates both the light direction and position from a parametric nonlinear regression model. Tasdizen et al. [60] developed an algorithm which is based on a semiparametric nonlinear regression model which is easily interpretable and flexible enough to take into account any smooth shape of illumination.

In this paper, we elaborate on the statistical model of Tasdizen et al. [60]. We propose a new reflectance estimation procedure which incorporates a new identifiability constraint. Our estimation procedure necessarily contains a denoising step. The statistical consistency of the reflectance estimate is studied through the criterion of the sup-norm risk. The study of the statistical consistency of the procedure takes into account both the impact of the denoising step and the impact of the illumination estimation error. A denoising method has then to be chosen to fulfil some conditions derived later. We show that these conditions are fulfilled when using multivariate local polynomials in the denoising step. We insist on the fact that any denoising method could be used in our procedure provided that both the image and its gradient are consistently recovered. It would be very interesting if we could manage to derive the rate of consistency for more recent and more elaborate denoising methods. This is left for future research.

We set the mathematical illumination–noise model in Section 2 and propose a semiparametric estimation procedure. In particular, we identify, discuss, and propose a solution to an identifiability issue which was still an open problem, up to our knowledge. In Section 3, we investigate the consistency properties of our estimation procedure. In particular, we derive upper bounds for the sup-norm risk of our estimation procedure. Practical choices of both the parametric and the nonparametric components of our estimation procedure are discussed in Section 4, as well as the suitability of multivariate local polynomials as a denoising method. Finally, applications to real digital images are presented in Section 5 together with a comparison experiment with different image processing techniques. Some technical details are postponed in the Appendix.