Available online at www.sciencedirect.com
SCIENCE DIRECT®

Pattern Recognition Letters 27 (2006) 1336–1341

Pattern Recognition Letters

www.elsevier.com/locate/patrec

An adaptive algorithm for image restoration using combined penalty functions

Daan Zhu *, Moe Razaz, Mark Fisher

School of Computing Sciences, University of East Anglia, Norwich NR4 7TJ, United Kingdom

Received 20 February 2005; received in revised form 8 January 2006
Available online 20 March 2006

Communicated by M.A.T. Figueiredo

Abstract

In this paper, we present an adaptive gradient based method to restore images degraded by the effects of both noise and blur. The approach combines two penalty functions. The first derivative of the Canny operator is employed as a roughness penalty function to improve the high frequency information content of the image and a smoothing penalty term is used to remove noise. An adaptive algorithm is used to select the roughness and smoothing control parameters. We evaluate our approach using the Richardson–Lucy EM algorithm as a benchmark. The results highlight some of the difficulties in restoring blurred images that are subject to noise and show that in this case an algorithm that uses a combined penalty function is able to produce better quality results.

Keywords: Gradient descent; Penalized likelihood; Combined penalty function; Image restoration; Regularization

1. Introduction and previous work

When an image passes through an optical system it appears blurred. This process can be modelled by convolving the image with the impulse response of the system. The image acquisition process also introduces noise which is often modelled as either a Gaussian or Poisson process depending on the type of image sensor and associated electronics (Pratt, 2003). Many linear and nonlinear algorithms have been developed in the literature to restore such noisy and blurred images, but both removing noise and sharpening the image remains a very challenging problem. A particular problem is that approaches aimed at restoring the effects of blur may in fact amplify the noise and introduce other unwanted image artifacts.

Constrained algorithms are quite a popular approach with many papers published concerning methods for blurred image restoration, which at the same time suppress the noise (see for example Carasso, 1999; Razaz et al., 1997). There are two difficulties in implementing a constrained equation. One is in selecting the penalty function and the other lies in finding the Lagrange parameter. Some recently published work has proposed various constrained maximum likelihood algorithms to reduce ringing effects (Lantéri et al., 2002a) or distortions introduced by iterative approaches (Heijden and Glasbey, 2003). A popular penalty term is a positivity constraint which is used to ensure that negative pixel values do not occur in the maximum solution that is in turn obtained by an expectation maximization (EM) algorithm (Lantéri et al., 2001; Lantéri et al., 2002b). Vogel and Oman (1998) approach restoration of noisy and blurred images by a Tikhonov regularization with a modified total variation function. In this paper, we focus on the problem of selecting a penalty function and finding an optimum Lagrange parameter.

A regularized equation can be expressed as (Katsaggelos and Kang, 1992)

$$\Gamma(\mu, \hat{f}) = \Gamma_1(\hat{f}) + \mu\Gamma_2(\hat{f}) \quad (1)$$

* Corresponding author. Fax: +44 0 1603 593345.
E-mail addresses: d.zhu@uea.ac.uk (D. Zhu), mr@cmp.uea.ac.uk (M. Razaz), mhf@cmp.uea.ac.uk (M. Fisher).

0167-8655/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2006.01.009

where Γ_1 and Γ_2 are the deconvolution and constraint parts, $\hat{f}$ is a guess image and μ a Lagrange parameter. A number of techniques have been presented aimed at selecting parameters. Razaz et al. (1997) developed a projection onto convex sets (POCS) method which uses a Newton–Raphson algorithm to estimate the Lagrange parameter. Recently, Chojnacki et al. (2004) discussed a constraint parameter estimator for image processing applications. Katsaggelos and Kang (1992) employed constraint least squares (CLS) to estimate the Lagrange parameter and Kang and Kastsaggelos (1995) implemented a function to find the noise control parameter. Thompson et al. (1991) used minimization of the total predicted mean squared error (TPMSE) and generalized cross-validatory (GCV) to calculate the Lagrange parameter. Ibáñez and Simó (2003) estimated a constraint parameter based on a Markov random field (MRF) model.

The above methods are based on the estimation of the smoothing term of a noise control parameter. However, in practice, we find that after introducing a smoothing parameter, noise can be removed but high frequency information in the image, due to edges, is blurred as well. Recently attempts to address this problem have focused on the selection of a roughness function. Typical approaches include Good's roughness regularization (Joshi and Miller, 1993), edge-preserving regularization (Machado et al., 2003) and k-means clustering (Charbonnier et al., 1997). In (Zhu et al., 2002) Zhu and Razaz present an adaptive algorithm based on the Canny filter to compensate for the loss of high frequency information. The Canny filter (Canny, 1986) is a well known algorithm for edge extraction which is robust to background noise in an image. In this work we employ Canny enhancement as the roughness parameter to enhance high frequency information.

A classical method for solving the likelihood equation is the Richardson–Lucy algorithm which uses expectation maximization (EM) (Katsaggelos and Lay, 1991). Although EM is an efficient method for solving an unconstrained likelihood equation the solution may not be optimally convergent when subject to constraints (Hebert and Lu Keming, 1995). In this paper, we use a gradient descent algorithm (GDA) to solve the penalized likelihood equation. This approach exhibits better convergence properties and noise removal performance than the Richardson–Lucy algorithm. Additionally, the GDA can also handle the problem of an image degraded by a spatially variant blur operator.

2. Mathematical modeling

The likelihood of image pixel intensity in an observed image, assuming a Poisson noise distribution, can be represented as

$$P(g|f) = \prod_i \left[\frac{(Hf)_i^{g_i}}{g_i!} \mathrm{e}^{-(Hf)_i}\right] \tag{2}$$

where i denotes the number of events in unit time, H the degradation matrix, g the noisy and blurred observed image and f the source image. The log-likelihood becomes

$$T(f) = \log(P(g|f)) = \log\left[\prod_i \left[\frac{(Hf)_i^{g_i}}{g_i!} \mathrm{e}^{-(Hf)_i}\right]\right] \tag{3}$$

then

$$T(f) = \sum g_i \log(Hf)_i - \sum \log(g_i!) - \sum (Hf)_i \tag{4}$$

Using Stirling's formula to expand $\log(g_i!)$ and dropping the independent terms f from $T(f)$ then

$$\Gamma(f) \approx \sum_{i=0}^{MN} [g_i - (Hf)_i] - \sum_{i=0}^{MN} g_i \frac{\log g_i}{(Hf)_i} \tag{5}$$

where M and N denote the width and height of the original image

$$\Gamma(f) = \sum_{i=0}^{MN} [g_i \log(Hf)_i - (Hf)_i] \tag{6}$$

Eq. (6) represents the formula for an unconstrained Γ function. By introducing a smoothing penalty term into Eq. (6), Hudson and Lee (1998) show that noise in the degraded image can be reduced, but high frequency components are also suppressed. We introduce both a roughness penalty term (Zhu et al., 2002) and a smoothing term giving

$$\Phi(\hat{f}) = \sum_{i=0}^{NM} [(Hf)_i - g_i \log(Hf)_i] + \alpha S(\hat{f}) + \beta C(\hat{f}) \tag{7}$$

where $S(\hat{f})$ and $C(\hat{f})$ represent the smoothing and roughness penalty terms. The minimum solution of $\Phi(\hat{f})$ is the optimal solution under constraints S and C.

2.1. Smoothing term selection

The penalty smoothing term, $S(\hat{f})$, which is used to suppress noise, is chosen to be equal to $\|R\hat{f}\|_2^2$, where R is the following diagonal matrix:

$$R = \begin{pmatrix} 1 & -0.75 & -0.25 & 0 & \cdots & 0 & 0 \\ -0.75 & 1 & -0.75 & -0.25 & \cdots & 0 & 0 \\ -0.25 & -0.75 & 1 & -0.75 & \cdots & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \cdots & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & -0.75 & -0.25 \\ 0 & 0 & \ddots & -0.25 & -0.75 & 1 & -0.75 \\ 0 & 0 & \ldots & 0 & -0.25 & -0.75 & 1 \end{pmatrix}$$

The selection of R depends on the noise properties. Due to the characteristics of the imaging system the noise in any one pixel is only related to the intensity in a small number of neighboring pixels. A more detailed discussion of coefficient matrix R is given in (Zhu et al., 2005; Hanke et al., 2000).