

Discrete Calculus, Optimisation and Inverse Problems in Imaging

Hugues Talbot^(✉)

Université Paris-Est ESIEE, Paris, France

`hugues.talbot@esiee.fr`, `hugues.talbot@univ-paris-est.fr`

Abstract. Inverse problems in science are the process of estimating the causal factors that produced a set of observations. Many image processing tasks can be cast as inverse problems: image restoration: noise reduction, deconvolution; segmentation, tomography, demosaicing, inpainting, and many others, are examples of such tasks. Typically, inverse problems are ill-posed, and solving these problems efficiently and effectively is a major, ongoing topic of research. While imaging is often thought of as occurring on regular grids, it is also useful to be able to solve these problems on arbitrary graphs. The combined frameworks of discrete calculus and modern optimisation allow us to formulate and provide solutions to many of these problems in an elegant way. This tutorial article summarizes and illustrates some of the research results of the last decade from this point of view. We provide illustrations and major references.

1 Introduction

Inverse problems are prevalent in science, because science is rooted in observations, and observations are nearly always indirect and certainly never perfect: noise is inevitable, instruments suffer from limited bandwidth, various artifacts and far from faultless acquisition components. Yet there is widespread interest in getting the most out of any observations one can make. Indeed, observations at the frontiers of science are typically those that are the faintest, most blurred, noisy and imperfect.

Imaging science is no exception, and this is why, for many decades [37], imaging communities have worked tirelessly to develop effective methods of noise removal, deconvolution, tomography, segmentation, inpainting, and more, in order to extract useful measurements from imperfect data.

Typically, inverse problems are ill-posed [26], meaning that their solution varies widely depending on the input data, in particular with noise. A useful approach for solving inverse problems is to use some sort of prior information on the data, called regularization. To formalize this, we turn to statistics.

2 Statistical Interpretation of Inverse Problems

We want to estimate some statistical parameter θ on the basis of some observation vector \mathbf{x} .

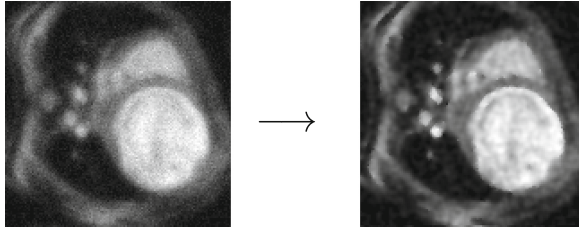


Fig. 1. An inverse problem: we observe the blurry, noisy data on the left. We would like to measure data from the image on the right.

2.1 Maximum Likelihood

If f is the sampling distribution, $f(x|y)$ is the probability of x when the population parameter is y . The function

$$y \mapsto f(x|y) \quad (1)$$

is called the *likelihood*. The Maximum Likelihood (ML) estimate is

$$\hat{y}_{\text{ML}}(x) = \underset{y}{\operatorname{argmax}} f(x|y) \quad (2)$$

A very common example, assuming we have a linear operator \mathbf{H} (in matrix form) and Gaussian deviates, then

$$\underset{y}{\operatorname{argmax}} f(y) = -\|\mathbf{H}y - x\|_2^2 = -y^\top \mathbf{H}^\top \mathbf{H} y + 2x^\top \mathbf{H} y - x^\top x \quad (3)$$

is a quadratic form with a unique maximum, provided by the *normal equations*:

$$\nabla f(y) = -2\mathbf{H}^\top \mathbf{H} y + 2\mathbf{H}^\top x = 0 \Rightarrow \hat{y} = (\mathbf{H}^\top \mathbf{H})^{-1} \mathbf{H}^\top x \quad (4)$$

This simple least-square formulation is very general, yet versions of the Maximum Likelihood solution correspond to a large class of problems in statistics and imaging, from simple linear regression to the Wiener filter [28,39] used in signal and image restoration [38], tomography with the filtered back-projection method [36], and many others. When it can be used, the ML solution is fast and effective. However the ML solution requires a descriptive model (with few degrees of freedom) and a lot of data, which is often unsuitable for images because we do not have a suitable model for natural images. When we do not have all these hypotheses, sometimes the Bayesian Maximum A Posteriori approach can be used instead.

2.2 Maximum a Posteriori

If we assume that we know a *prior* distribution g over y , i.e. some *a-priori* information. Following Bayesian statistics, we can treat y as a random variable and compute the *posterior* distribution of y , via the Bayes theorem:

$$y \mapsto f(y|x) = \frac{f(x|y)g(y)}{\int_{\vartheta \in \Theta} f(x|\vartheta)g(\vartheta)d\vartheta} \quad (5)$$

Then the Maximum a Posteriori is the estimate

$$\hat{y}_{\text{MAP}}(x) = \underset{y}{\operatorname{argmax}} f(y|x) = \underset{y}{\operatorname{argmax}} f(x|y)g(y) \quad (6)$$

We say that the MAP estimate is a *regularization* of ML. The only difference between ML and MAP is the $g(y)$ multiplicative terms. For easier handling, we take the log, which does not change the estimator since the log function is monotonic:

$$\hat{y}_{\text{MAP}}(x) = \underset{y}{\operatorname{argmax}} \log f(x|y) + \log g(y). \quad (7)$$

The first term of the right-hand side is call the log-likelihood, and the second term is the regularisation. In optimisation theory, a minimization is usually preferred, so we simply multiply (7) by -1 . In particular, the log-likelihood become the negative log-likelihood [4, chap. 7].

3 Imaging Formulations

The very brief exposition of the previous section covers the basic principle of many statistical methods, including PCA, LDA, EM, Markov Random Fields, Hidden Markov Models, up to graph-cut type methods in imaging [5]. Many details can be found in classical texts on pattern analysis [22].

In the case of imaging, the log-likelihood term is often called the *data fidelity*. If we assume an image $\bar{y} \in \mathbb{R}^N$ is corrupted by noise and blur, for instance, we observe the data $x \in \mathbb{R}^Q$, and we can write:

$$x = \mathbf{H}\bar{y} + u, \quad \mathbf{H} \in \mathbb{R}^{Q \times N}, \quad (8)$$

where u is the noise. \mathbf{H} can typically be a camera model including blur and defocus, a tomography projection matrix, an MRI analysis matrix (i.e. a Fourier transform), etc. The noise model is here additive, but with some work it is possible to express the likelihood of more complex noises: Rician, Poisson, Poisson-Gauss [27], etc. To recover an estimation \hat{y} of \bar{y} , the ML estimator is, in the additive Gaussian noise case, the least-square estimate:

$$\hat{y} = \underset{y}{\operatorname{argmin}} \|\mathbf{H}y - x\|_2^2. \quad (9)$$

Often this is not robust. A number of MAP regularizations can be proposed. The simplest is the Tikhonov regularisation:

$$\hat{y} = \underset{y}{\operatorname{argmin}} \|\mathbf{\Gamma}y\|_2^2 + \lambda \|\mathbf{H}y - x\|_2^2, \quad (10)$$

where λ is a Lagrangian multiplier, and $\mathbf{\Gamma}$ is a linear operator, which can be the identity or a spatial gradient for instance. The corresponding quadratic prior

term expresses the belief that $\Gamma\mathbf{y}$ has a zero-centered Gaussian distribution, i.e. is typically smooth. This model is very easy to optimize but not realistic for most images, although it can be related to anisotropic diffusion for denoising [3, 33], and the Random Walker for segmentation [24].

A more interesting approach for imaging is to define a *sparsity* prior. If we assume \mathbf{y} to be sparsely representable, for instance in a wavelet basis, then it might be interesting to use this in a regularization prior. Ideally, one would like to use the ℓ_0 pseudo-norm to enforce sparsity, However this pseudo-norm is both non convex and non-differentiable, which makes it difficult to use in practice. A key element of *compressive sensing* [7] is based on the observation that the ℓ_1 norm is nearly as effective at promoting sparsity.

$$\hat{\mathbf{y}} = \underset{\mathbf{y}}{\operatorname{argmin}} \|\Gamma\mathbf{y}\|_1 + \lambda\|\mathbf{H}\mathbf{y} - \mathbf{x}\|_2^2. \quad (11)$$

We will now explore some of these priors. Before we can do that, we need to propose a way to define flexible-enough linear operators well suited to imaging.

4 Linear Operators

The classical operators in continuous-domain formulations of the problems we have seen so far are the gradient and its adjoint the divergence. These can be easily discretized using finite-difference schemes [20]. Continuous and discrete versions of wavelet operators can also be considered. In the sequel, we choose to define our operators on arbitrary graphs, in the framework of discrete calculus [25].

4.1 The Incidence Matrix

Given a directed graph of N vertices and M edges, we can define the $M \times N$ incidence matrix \mathbf{A} , with lines containing zeros and exactly one $+1$ and one -1 so that $a_{i,k} = -1$ and $a_{i,l} = +1$ if e_i is the (k, l) edge. An illustrative example is best at this point (see Fig. 3). The matrix \mathbf{A} describes the graph but can also be thought of as an operator. If \mathbf{p} is a vector of values associated to the vertices, then $\mathbf{A}\mathbf{p}$ is the gradient operator, associating a value to every edge. The transpose matrix \mathbf{A}^\top is the adjoint operator, corresponding to the negative divergence (Fig. 2).

4.2 The Dual-Constrained Total Variation Model

Among the interesting regularizations, the Total Variation (TV) [35], or ROF model after the initials of its inventors, promotes sparsity of the gradient. In other words, it corresponds to a piecewise-constant image prior. This is of interest for various reasons, one of which because it is an acceptable model for texture-free natural images. Simplified versions of the Mumford-Shah model [30] for image segmentation typically use a TV term instead of the more complex piecewise-smooth prior. In [8], authors introduce TV formulations for image restoration in a MAP framework.

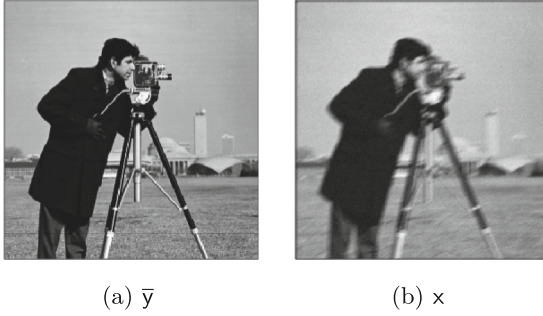


Fig. 2. Ideal and observed image

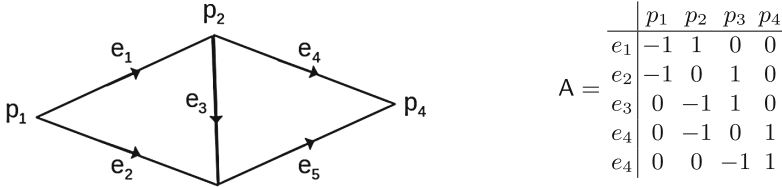


Fig. 3. A small graph and its associated incidence matrix.

A weighted version of the TV model can be written in the following way [23], in the continuous framework:

$$\min_u \underbrace{\int \left(\int w_{x,y} (u_y - u_x)^2 dy \right)^{1/2} dx}_{\text{regularization } R(u)} + \underbrace{\frac{1}{2\lambda} \int (u_x - f_x)^2 dx}_{\text{data fidelity } \Phi(u)}, \quad (12)$$

with λ a Lagrange multiplier. It is equivalent to the following min-max problem [10]

$$\min_u \max_{\|p\|_\infty \leq 1} \int \int w_{x,y}^{1/2} (u_y - u_x) p_{x,y} dx dy + \Phi(u), \quad (13)$$

with p a projection vector field. Such min-max formulations are called primal-dual in optimization. The field p is introduced to achieve better speed. Constraining p can promote better results, as we will see in Sect. 6.

In discrete calculus form, we can write the same problem in this way:

$$\min_u \max_{\|p\|_\infty \leq 1, p \in \mathbb{R}^M} p^\top ((\mathbf{A}u) \cdot \sqrt{\mathbf{w}}) + \Phi(u). \quad (14)$$

Introducing the projection vector $\mathbf{F} \in Rset^M = p \cdot \sqrt{\mathbf{w}}$, we can constrain \mathbf{F} to belong to a convex set $C = \bigcap_{i=1}^{m-1} C_i \neq \emptyset$ where C_1, \dots, C_{m-1} closed convex sets of \mathbb{R}^M . Given $\mathbf{g} \in \mathbb{R}^N$, $\theta_i \in \mathbb{R}^M$, $\alpha \geq 1$, $C_i = \{\mathbf{F} \in \mathbb{R}^M \mid \|\theta_i \cdot \mathbf{F}\|_\alpha \leq g_i\}$,

$$\min_{u \in \mathbb{R}^N} \sup_{\mathbf{F} \in C} \mathbf{F}^\top (\mathbf{A}u) + \Phi(u). \quad (15)$$

The \mathbf{F} constraints can be interpreted as *flow* constraints on the vertices of the connecting graph. For image denoising, we can for example propose that $g_i \in \mathbb{R}^N$ be a weight on vertex i , inversely function of the gradient of f at node i . In this case:

- Over flat areas: weak gradient implies a strong g_i , itself implying a strong $F_{i,j} \rightarrow$ weak local variations of u .
- Near contours: strong gradient implies a weak g_i itself implying a weak $F_{i,j} \rightarrow$ large local variations of u are allowed.

This model is the dual-constrained total variation (DCTV) [17]. To optimize it, we require algorithms capable of dealing with non-differentiable convex functionals.

5 Algorithms

Optimization algorithms are numerous but research have mostly focused on differentiable methods: gradient descent, conjugate gradient, etc [4], with the exception of the simplex method for linear programming [21]. However non-differentiable optimization methods have been available at least since the 1960s. The main tool for non-differentiable optimizing convex functionals is the proximity operator [1, 14, 29, 34]. We recall here the main points.

5.1 Proximity Operator

Let $\Gamma_0(\mathbb{R}^N)$ be the set of proper (i.e. not everywhere equal to $+\infty$), lower semi-continuous, convex functionals taking values from \mathbb{R}^N to $(-\infty, +\infty]$. Such functions are necessarily quite regular. In particular, they are continuous and almost everywhere differentiable. The subdifferential of $f \in \Gamma_0(\mathbb{R}^N)$ at \mathbf{x} is given by

$$\partial f(\mathbf{x}) = \{\mathbf{u}, \forall \mathbf{y} \in \mathbb{R}^N, (\mathbf{y} - \mathbf{x})^\top \mathbf{u} + f(\mathbf{x}) \leq f(\mathbf{y})\} \tag{16}$$

This definition extends the notions of tangent and thus of derivative to the non-differentiable case. Where f is differentiable, the subdifferential and the derivative are equal. We note that the subdifferential at non-differentiable points is a set, not a scalar or a vector.

The proximity operator of f in \mathbf{x} is the operator $\mathbb{R}^N \rightarrow \mathbb{R}^N$

$$\text{prox}_f(\mathbf{x}) = \underset{\mathbf{y} \in \text{dom}(f)}{\text{argmin}} f(\mathbf{y}) + \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 \tag{17}$$

We have the following property

$$\mathbf{p} = \text{prox}_f(\mathbf{x}) \Leftrightarrow \mathbf{x} - \mathbf{p} \in \partial f(\mathbf{p}), \forall (\mathbf{x}, \mathbf{p}) \in \mathbb{R}^N \times \mathbb{R}^N \tag{18}$$

5.2 Splitting

One of the simplest cases is the situation when one wants to optimize the sum of two functions, one of which is smooth. Let $f_1 \in \Gamma_0(\mathbb{R}^N)$ and $f_2 : \mathbb{R}^N \rightarrow \mathbb{R}$ convex and differentiable with a β -Lipschitz constant gradient ∇f_2 , i.e.

$$\forall(x, y) \in \mathbb{R}^N \times \mathbb{R}^N, \|\nabla f_2(x) - \nabla f_2(y)\| \leq \beta\|x - y\|, \quad (19)$$

with $\beta > 0$. If $f_1(x) + f_2(x) \rightarrow +\infty$ when $\|x\| \rightarrow +\infty$ (i.e. $f_1 + f_2$ is coercive), we wish to

$$\underset{x \in \mathbb{R}^N}{\text{minimize}} \quad f_1(x) + f_2(x). \quad (20)$$

It can be shown that this problem admits a solution and that for any $\gamma > 0$, the following fixed-point equation holds

$$x = \text{prox}_{\gamma f_1}(x - \gamma \nabla f_2(x)). \quad (21)$$

This suggests the following explicit-implicit algorithm

$$x_{n+1} = \text{prox}_{\gamma f_1}(x_n - \gamma \nabla f_2(x_n)). \quad (22)$$

This algorithm is the forward-backward, alternating an explicit forward gradient descent step with an implicit proximity operator backward step. It can be shown [15] that this algorithm converges to a solution to (20).

This fairly simple method extends well-known ones such as gradient descent and the proximal point algorithm. It can be improved, for instance replacing the gradient descent scheme with Nesterov's method [32], which in this case yields an optimal convergence rate [31]. The corresponding method is the Beck-Teboule proximal gradient method [2].

5.3 Primal-Dual Methods

Many splitting methods exist, involving sums of two or more functions, and are detailed in [14]. In the case of (15), the presence of explicit constraints makes the analysis more difficult. Using convex analysis, and in particular the Fenchel-Rockafeller duality, and if the graph is regular, we can optimize it using the Parallel Proximal Algorithm (PPXA) [13]. In the more interesting case of an irregular graph, a primal-dual method is necessary [9]. We actually used the algorithm detailed in [6], which has since been generalized [12, 16].

We now show some results obtained from solving (15) in various contexts.

6 Results

DCTV is a flexible framework. In Fig. 4(a,b,c) we restore a blurry, noisy version of an MRI scan using a local regular graph. This is the same image as in Fig. 1. In Fig. 4(d,e,f) we restore an image using an irregular, non-local graph. The fine texture of the brick wall has been restored to a high degree. In Fig. 4(g,h,i) we restore a noisy 3D mesh, with the same framework. Only the graph changes.

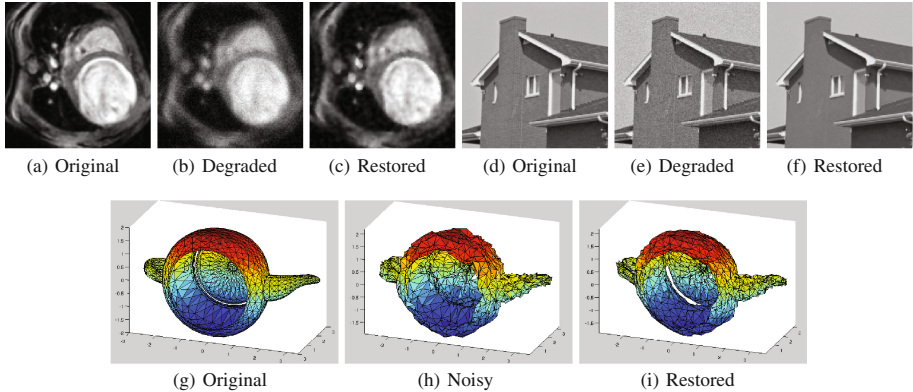


Fig. 4. (a) Original MRI image, (b) Noisy, blurred image $\text{SNR} = 12.3\text{ dB}$, (c) DCTV result $\text{SNR} = 17.2\text{ dB}$. (d) Original house image, (e) Noisy image $\text{PSNR} = 28.1\text{ dB}$, (f) DCTV result $\text{PSNR} = 35\text{ dB}$. (g) Original mesh, (h) noisy mesh, (i) restored mesh

7 Discussion

Results presented here are interesting to some degree because we have kept the spatial part of the formulation fully discrete, with at its heart a graph representation for the numerical operators we use. However an important point is our assumption that the distribution of image values is continuous. In practice this is not the case and our approach is a relaxation of the reality, since images are typically discretized to 8 or 16-bit values. If we require to keep discretized values throughout the formulation, for instance to deal with labeled images, then the approach proposed here would not work. In this case, MRF formulations could be used [18, 19].

We have also kept the discussion in the convex framework. Many important problems are not convex, for instance blind deblurring, where the degradation kernel must be estimated at the same time as the restoration. There exist methods for dealing with non-convex image restoration problems, for instance [11], but dealing with non-convexity and non-differentiability together remains a challenge in the general case.

8 Conclusion

In this short overview article, we have introduced inverse problems in imaging and a statistical interpretation: the MAP principle for solving inverse problems such as image restoration using a-priori information. We have shown how we can use a graph formulation of numerical operators using discrete calculus to propose a general framework for image restoration. This DCTV framework can be optimized using non-differentiable convex optimization techniques, in particular the proximity operator. We have illustrated this approach on several examples.

DCTV is by no means the only framework available but it is one of the most flexible, fast and effective. With small changes we can tackle very different problems such as mesh or point cloud regularization. In general, the combination of powerful optimization methods, graph representations of spatial information and fast algorithms is a compelling approach for many applications.

References

1. Bauschke, H., Combettes, P.: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York (2011)
2. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2**(1), 183–202 (2009)
3. Black, M.J., Sapiro, G., Marimont, D.H., Heeger, D.: Robust anisotropic diffusion. *IEEE Trans. Image Process.* **7**(3), 421–432 (1998)
4. Boyd, S., Vandenberghe, L.: *Convex Optimization*. Cambridge University Press, New York (2004)
5. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(11), 1222–1239 (2001)
6. Briceno-Arias, L.M., Combettes, P.L.: A monotone+ skew splitting model for composite monotone inclusions in duality. *SIAM J. Optim.* **21**(4), 1230–1250 (2011)
7. Candes, E.J., Romberg, J.K., Tao, T.: Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.* **59**(8), 1207–1223 (2006)
8. Chambolle, A., Caselles, V., Cremers, D., Novaga, M., Pock, T.: An introduction to total variation for image analysis. *Theor. Found. Numer. Methods Sparse Recovery* **9**(263–340), 227 (2010)
9. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.* **40**(1), 120–145 (2011)
10. Chan, T.F., Golub, G.H., Mulet, P.: A nonlinear primal-dual method for total variation based image restoration. *SIAM J. Sci. Comput.* **20**(6), 1964–1977 (1999)
11. Chouzenoux, E., Jezierska, A., Pesquet, J.-C., Talbot, H.: A majorize-minimize subspace approach for ℓ_2 - ℓ_0 image regularization. *SIAM J. Imaging Sci.* **6**(1), 563–591 (2013)
12. Combettes, P., Pesquet, J.: Primal-dual splitting algorithm for solving inclusions with mixtures of composite, lipschitzian, and parallel-sum type monotone operators. *Set-Valued Variational Anal.* **20**(2), 307–330 (2012)
13. Combettes, P.L., Pesquet, J.-C.: A proximal decomposition method for solving convex variational inverse problems. *Inverse Prob.* **24**(6), 065014 (2008)
14. Combettes, P.L., Pesquet, J.-C.: Proximal splitting methods in signal processing. In: Bauschke, H.H., Burachik, R., Combettes, P.L., Elser, V., Luke, D.R., Wolkowicz, H. (eds.) *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. Springer, New York (2010)
15. Combettes, P.L., Wajs, V.R.: Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.* **4**(4), 1168–1200 (2005)
16. Condat, L.: A primal-dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms. *J. Optim. Theory Appl.* **158**(2), 460–479 (2013)
17. Couprie, C., Grady, L., Najman, L., Pesquet, J.-C., Talbot, H.: Dual constrained TV-based regularization on graphs. *SIAM J. Imaging Sci.* **6**(3), 1246–1273 (2013)

18. Couprie, C., Grady, L., Najman, L., Talbot, H.: Power watersheds: a new image segmentation framework extending graph cuts, random walker and optimal spanning forest. In: Proceedings of ICCV, Kyoto, Japan, 2009, pp. 731–738. IEEE (2009)
19. Couprie, C., Grady, L., Najman, L., Talbot, H.: Power watersheds: a unifying graph-based optimization framework. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(7), 1384–1399 (2011)
20. Courant, R., Friedrichs, K., Lewy, H.: On the partial difference equations of mathematical physics. *IBM J.* **11**(2), 215–234 (1967)
21. Dantzig, G.B.: *Linear Programming and Extensions*, 11th edn. Princeton University Press, Princeton (1998)
22. Duda, R.O., Hart, P.E., Stork, D.G.: *Pattern Classification*, 2nd edn. John Wiley & Sons, New York (2001)
23. Gilboa, G., Osher, S.: Nonlocal linear image regularization and supervised segmentation. *SIAM J. Multiscale Model. Simul.* **6**(2), 595–630 (2007)
24. Grady, L.: Random walks for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(11), 1768–1783 (2006)
25. Grady, L., Polimeni, J.: *Discrete Calculus: Applied Analysis on Graphs for Computational Science*. Springer Publishing Company, London (2010)
26. Hadamard, J.: Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton university bulletin* **13**(49–52), 28 (1902)
27. Jezierska, A., Chouzenoux, E., Pesquet, J.-C., Talbot, H., et al.: A convex approach for image restoration with exact poisson-gaussian likelihood. *IEEE Trans. Image Process.* **22**(2), 828 (2013)
28. Kolmogorov, A.: Stationary sequences in hilbert space. In: *Linear Least-Squares Estimation*, p. 66 (1941)
29. Moreau, J.-J.: Fonctions convexes duales et points proximaux dans un espace hilbertien. *CR Acad. Sci. Paris Sér. A Math.* **255**, 2897–2899 (1962)
30. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. *Commun. Pure Appl. Math.* **42**(5), 577–685 (1989)
31. Nemirovsky, A.-S., Yudin, D.-B., Dawson, E.-R.: *Problem Complexity and Method Efficiency in Optimization*. John Wiley & Sons Ltd, New York (1982)
32. Nesterov, Y.: Smooth minimization of non-smooth functions. *Math. Program.* **103**(1), 127–152 (2005)
33. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **12**(7), 629–639 (1990)
34. Rockafellar, R.T.: *Convex Analysis*. Princeton University Press, New Jersey (1970). Reprinted 1997
35. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Phys. D* **60**(1–4), 259–268 (1992)
36. Smith, S.W.: *The Scientist and Engineer’s Guide to Digital Signal Processing*. California Technical Publishing, San Diego (1999)
37. Twomey, S.: *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements*. Elsevier, New York (1977)
38. Widrow, B., Stearns, S.D.: *Adaptive Signal Processing*. Prentice-Hall Inc, Englewood Cliffs (1985)
39. Wiener, N.: *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*, vol. 2. MIT press, Cambridge (1949)