

Saddle point formulation for a cartoon-texture decomposition*

Stefan Kindermann[†] Stanley Osher[‡]

Abstract

We consider the image-decomposition into a cartoon and texture part proposed by Yves Meyer. We formulate the corresponding optimization problem as a saddle point problem and show the existence of a saddle point. This also leads to a dual problem from which we can find conditions on the regularization parameter when a trivial decomposition occurs. The saddle point formulation allows us to state the Euler-Lagrange equations for optimality, which take the form of a nonlinear indefinite system. We propose to choose the dual parameter as regularization parameter. Finally, we discuss some numerical algorithms for solving the optimality conditions and investigate their convergence numerically.

1 Introduction

The task of image decomposition is to split a given image f into a sum of two terms $f = u + v$ where u and v represent different information in the image. For instance, in the process of denoising a successful split of a noisy image f should result in a clean image u and a pure noisy part v . A closely related problem is to decompose an image into a 'cartoon' part u which is a sketch of the original image and a part v which contains the textures.

A useful framework to obtain a decomposition is the variational approach, where u is found by minimizing a certain functional. In general, such functionals are made of two terms: a fidelity term measuring the distance to the original image f and a regularization term, which penalizes high frequencies.

One of the best-known nonlinear method in this class is the Rudin-Osher-Fatemi functional (ROF) [19]: For a given grey scale image $f \in L^2$ the cartoon part u is defined as minimizer of

$$J(u) := \lambda \|u - f\|_{L^2}^2 + |u|_{BV}. \quad (1)$$

*Research supported by NIH U54RR021813, NSF DMS-0312222, NSF ACI-0321917 and NSF DMI-0327077

[†]Department of Mathematics, UCLA, 520 Portola Plaza, Los Angeles, Ca 90095, on leave from University Linz, Austria, email: kindermann@indmath.uni-linz.ac.at

[‡]Department of Mathematics, UCLA, 520 Portola Plaza, Los Angeles, Ca, 90095, email: sjo@math.ucla.edu

Here $\lambda > 0$ is a regularization parameter, which determines the regularity of the outcome u , and $|u|_{BV}$ is the bounded variation seminorm (2). The ROF-functional has been generalized in many ways by using different norms in (1) (see e.g. [8, 18]).

In this work we consider the generalization of the ROF-functional proposed by Y. Meyer [14]. The idea is to obtain a better image decomposition into a cartoon and texture part by replacing the L^2 -norm in (1) with a norm suitable for textures. Note that $v = f - u$ is the texture component of the decomposition. In [14] the author proposed the use of the G-norm as a reasonable norm for oscillating patterns such as textures, which leads to the functional (11). However, the analysis and computation of the related optimization problem is much more involved compared to the ROF-model.

The aim of this paper is to analyze the Meyer functional (11) theoretically and numerically. We will reformulate the problem of minimizing this functional into an equivalent saddle point problem. We will prove the existence of a saddle point and hence also the existence of a minimizer for (11). The saddle point formulation also gives rise to a dual problem, which can be used to derive some properties of a minimizer. Moreover, we derive the optimality condition for the saddle point problem in Euler-Lagrange form which leads to a nonlinear indefinite system of partial differential equations. For the numerical computation we propose some iterative algorithms and the use of the dual parameter μ instead of the usual choice λ as tuning parameter. In the final part we perform some numerical experiments comparing different algorithms to solve the optimality conditions.

The paper is organized as follows: In Section 2 we introduce the G-norm and state some of its properties. In Section 3 we define the Meyer-functional and examine the corresponding saddle point problem. In Section 4 we discuss some numerical schemes and test their performance.

2 G-norm

The G-norm has been introduced in [14] as an appropriate norm for textures in images. While in a ROF-decomposition the cartoon part u is related to the BV -norm it is considered that the right norm for the texture part is the G-norm, which we will define in the following. In order to have useful embedding properties we will restrict ourselves to dimensions less than 3, although definition (3) is valid in any dimension. Let us assume that $\Omega \subset \mathbb{R}^n$, with $n = 1$ or $n = 2$, is a simply connected domain with Lipschitz boundary. Denote by BV the space of functions of bounded variations

$$BV := \{u \in L^1(\Omega) \mid \|u\|_{BV} := \|u\|_{L^1} + |u|_{BV} < \infty\},$$

with the BV -seminorm ($\nabla \cdot$ denotes the divergence operator)

$$|u|_{BV} := \sup_{\phi \in [C_0^\infty(\Omega)]^2, \|\phi\|_{L^\infty} \leq 1} \int_{\Omega} u(x) \nabla \cdot \phi dx. \quad (2)$$

We start with the definition for the case $\Omega = \mathbb{R}^n$ following [14]: Let $f \in L^2(\mathbb{R}^n)$, then the G-norm $\|f\|_*$ is defined as

$$\|f\|_* := \inf\{\|g\|_{L^\infty(\mathbb{R}^n)} \mid \nabla \cdot g = f\}. \quad (3)$$

For Ω being a bounded domain, this definition has been generalized in [3]. It turns out, that in order to develop a similar theory as for $\Omega = \mathbb{R}^n$ some additionally boundary conditions for g are needed in this case. Denote by n the unit outward normal on $\partial\Omega$, then the G-norm on a bounded domain is defined by

$$\|f\|_* := \inf\{\|g\|_{L^\infty} \mid \nabla \cdot g = f, g \cdot n = 0 \text{ on } \partial\Omega\}. \quad (4)$$

If for a given function f no vector field g exists with the properties in (4), then we set $\|f\|_* := \infty$. However, this pathological case is not very important, as any function f in $L^2(\Omega)$ can be redefined to have finite G-norm by subtracting a constant. In fact it was proven in [3], that for a bounded domain the function with finite G-norm are precisely those which have zero mean:

$$\|f\|_* < \infty \Leftrightarrow \int_{\Omega} f(x) dx = 0 \quad \forall f \in L^2(\Omega).$$

In the following we always assume that for bounded domains this normalization condition holds:

$$\int_{\Omega} f(x) dx = 0. \quad (5)$$

Of course, this condition is not necessary if $\Omega = \mathbb{R}^n$.

The previous definitions are not very well suited for practical purposes, since it is rather difficult to minimize over all vector fields with $\nabla \cdot g = f$. In this paper we use the G-norm with a different but equivalent definition. In [13] it was shown that definitions (3) or (4), respectively, are equivalent to

$$\|f\|_* = \sup_{u \in BV} \frac{\int_{\Omega} f(x) u(x) dx}{|u|_{BV}}. \quad (6)$$

Moreover, if f is in $L^2(\Omega)$ and additionally satisfies condition (5) in the case of a bounded domain Ω the supremum in (6) is attained by a function $u \in BV$. This definition has the advantage that it avoids the computation of a vector field g .

The following fundamental inequality is extremely useful in the analysis of the G-norm: For $f \in L^2(\Omega)$ and any $u \in BV$ we have

$$\int_{\Omega} f(x) u(x) dx \leq |u|_{BV} \|f\|_* \quad (7)$$

which simply follows from (6). An important case occurs when for a pair of functions the previous inequality is an equality. Y. Meyer calls such a pair an extremal pair. Let us cite the definition from [14]:

Definition 2.1. Two functions $u \in BV(\Omega)$, $f \in L^2(\Omega)$ are called an extremal pair (u, f) if

$$\int_{\Omega} f(x)u(x)dx = |u|_{BV}\|f\|_*. \quad (8)$$

It is trivial to show that for a given f u is a maximizer in (6) if and only if (u, f) is an extremal pair. It is known that an extremal pair is related by a partial differential equation (see [13, 14]). In fact, (u, f) is an extremal pair if and only if

$$\frac{f}{\|f\|_*} = -\nabla \cdot \frac{\nabla u}{|\nabla u|}. \quad (9)$$

The function u is not uniquely determined by conditions (8) or (9) because a multiplication by a nonzero scalar λu gives another extremal pair. Moreover, it was shown in [13] that if (u, f) is an extremal pair, then almost all the upper level set $\chi_{\tau} := H(u - \tau)$ of u also form extremal pairs (χ_{τ}, f) . (Here H denotes the Heaviside function). It should also be noted that the notion of extremal pairs is not symmetric, if (u, f) is one then (f, u) does not have to be an extremal pair.

By the G-norm and the notion of extremal pair a complete characterization of the ROF-decomposition can be given. The following theorem was proven in [14]:

Theorem 2.2. Let $f \in L^2$ and $u \in BV$ be nonzero. Then u is a minimizer of (1), if and only if $(u, f - u)$ is an extremal pair.

There are very few examples of functions f for which the G-norm can be computed analytically. For instance, when f is the characteristic function of a circle with radius R , then its G-norm is $\frac{R}{2}$. If f is the characteristic function of a square, it can be proven that the G-norm is proportional to the length of the square. These cases can be computed, since functions are known which form an extremal pair with f . For the characteristic function of a circle f makes an extremal pair with itself [14], while for the square the corresponding function is a rounded square [13, 20].

Finally we state some useful embedding theorems. If the dimension n is less than 3 then it is well known that the space BV can be embedded into L^2 ([1]). By (6) it is easy to show that the space L^2 (with the normalization condition (5)) can be continuously embedded into the space of functions with finite G-norm. Hence for functions f satisfying (5) we have the following embedding inequality:

$$\|f\|_* \leq C_1 \|f\|_{L^2} \leq C_2 |f|_{BV}. \quad (10)$$

This inequality was proven in [14]. For the case $\Omega = \mathbb{R}^2$ the embedding constants can be taken as $C_1 = \frac{1}{2\sqrt{\pi}}$ and $C_2 = \frac{1}{4\pi}$ ([11]).

3 Meyer-decomposition

This image decomposition was introduced by Yves Meyer in [14] as a variant of the ROF-decomposition. The main difference is that the G-norm is used instead

of the L^2 -norm to measure the texture part. The motivation for this lies in the fact that the v -component should contain textures, hence it seems reasonable to measure the residual part $v = f - u$ by the G-norm, which is considered a good norm for textures. Similar as for the ROF-model this leads to a variational problem. In fact, the Meyer-decomposition $f = u + v$ defines the cartoon-part u as a minimizer of the following functional

$$M(u) := \lambda \|f - u\|_* + |u|_{BV}, \quad (11)$$

where λ is the regularization parameter.

The use of the G-norm should give a better decomposition of an image into texture and cartoon part. However, it is more difficult to handle theoretically and numerically. Existence of a minimizer has been proven in [4] for the discrete case. (For the continuous case this will also follow from the saddle point formulation below in Theorem 3.3). Differently to the ROF-functional the minimizers are not unique in general. An example with a non unique decomposition is the case when f is the characteristic function of a disc, as was shown in [11].

Let us remark that in the case of Ω being bounded we can without loss of generality assume that $\int_{\Omega} f(x)dx = \int_{\Omega} u(x)dx = 0$. From the definition of the G-norm it follows that $M(u) < \infty$ implies $\int_{\Omega} f(x) = \int_{\Omega} u(x)$. If f does not have zero mean we can redefine $\tilde{f} := f(x) - \int_{\Omega} f(x)dx$, and $\tilde{u} := u(x) - \int_{\Omega} f(x)dx$. Then u is minimizer of $M(u)$ with f if and only if \tilde{u} is a minimizer of (11) with \tilde{f} in place of f . Moreover, a necessary condition for \tilde{u} is then $\int_{\Omega} \tilde{u}(x)dx = 0$.

Let us briefly discuss some related models. The main difficulty in this variational problem is the computation of the G-norm. Several variants of the functional (11) have been proposed. Vese and Osher [21, 22], replaced the L^∞ -norm in Definition by an L^p -norm with $p < \infty$. They used the the following functional

$$F_{VO}(u, g) = |u|_{BV} + \lambda \|f - u - \nabla \cdot g\|_{L^2}^2 + \mu \|g\|_{L^p}. \quad (12)$$

Here λ, μ are real parameter, u plays the role of the cartoon part, g is a vector field, and $p \in \mathbb{N}$. The cartoon-part for this image decomposition is found by a joint minimization over u and the vector field g . The use of the L^p -norm with $p < \infty$ instead of the L^∞ -norm makes it easier to deduce the Euler-Lagrange-equation and also easier to compute a numerical solution.

Another variant was proposed by Osher, Solè and Vese in [15]: The idea is to assume that the vector field g is a gradient field: $g = \nabla P$. Motivated by the definition of the G-norm the authors proposed to minimize the functional

$$J_{OSV}(u) := |u|_{BV} + \lambda \|\nabla \Delta^{-1}(f - u)\|_{L^2}, \quad (13)$$

where Δ is the Laplace-operator. In contrast to (12), this is a one-parameter model with a minimization only over u .

Another approach was suggested by Aujol et al. [4]. Their model is similar to the Vese-Osher version, but they use the full G-norm. Similar as (12) it is a model with two tuning parameter λ, μ , hence it gives a decomposition into three

parts $f = u + v + w$. Here u is regarded as the cartoon part, v the texture part and w the noise part. The authors in [4] propose to minimize the functional

$$FAABC(u, v) = |u|_{BV} + \lambda \|f - u - v\|_{L^2}^2 + \mu J^*(v), \quad (14)$$

where

$$J^*(v) := \begin{cases} 0 & \text{if } \|v\|_* \leq 1 \\ \infty & \text{else} \end{cases}$$

The u -part in an image decomposition is computed by minimizing over u and v . A minimizer of this functional can be computed by performing two ROF-decompositions simultaneously. We will describe this algorithm below. Moreover, it was shown that for $\lambda \rightarrow \infty$ the minimizers of (14) are also minimizers of (11).

Finally let us mention that for the discretized case a method based on second-order cone programming was proposed in [23].

Our starting point comes directly from the functional (11). We will reformulate the problem as an inf-sup problem, which allows us to apply methods of convex analysis to the problem. But in contrast to (12) and (14) we do not get a threefold decomposition $f = u + v + w$, since we only consider a one-parameter model.

Using definition (6) we can reformulate the optimization problem (11) into

$$\inf_{u \in BV} M(u) = \inf_{u \in BV} \sup_{p \in BV, |p|_{BV} \leq 1} L(u, p)$$

with

$$L(u, p) = \lambda \int_{\Omega} (f(x) - u(x)) p(x) dx + |u|_{BV}. \quad (15)$$

It is clear from definition (6) that \bar{u} is a minimizer of (11) if and only if it also satisfies

$$\bar{u} = \operatorname{argmin}_{u \in BV} \sup_{p \in BV, |p|_{BV} \leq 1} L(u, p).$$

The formulation as a inf-sup problem allows us to look at the dual problem, which is obtained by swapping the infimum and supremum:

Definition 3.1. *The dual problem to (11) is to find a $\bar{p} \in BV$, $|\bar{p}|_{BV} \leq 1$, solution to the following optimization problem*

$$G(\bar{p}) = \sup_{p \in BV, |p|_{BV} \leq 1} G(p),$$

where

$$G(p) := \inf_{u \in BV} L(u, p). \quad (16)$$

The following lemma shows, that the variable u in (16) can be completely eliminated, hence, the dual problem can be rewritten as a simple constraint optimization problem:

Lemma 3.2. *Let $f \in L^2(\Omega)$, $\lambda > 0$, then*

$$G(p) = \begin{cases} \lambda \int_{\Omega} f(x)p(x)dx & \text{if } \|p\|_* \leq \frac{1}{\lambda} \\ -\infty & \text{if } \|p\|_* > \frac{1}{\lambda} \end{cases}. \quad (17)$$

\bar{p} is a solution to the dual problem if and only if it is a solution to the constraint problem

$$\bar{p} = \operatorname{argmax}_{\|p\|_* \leq \frac{1}{\lambda}, |p|_{BV} \leq 1} \lambda \int_{\Omega} p(x)f(x)dx. \quad (18)$$

Moreover, under the given assumptions on f and λ this optimization problem always has a solution \bar{p} .

Proof. First we show equation (17). For $p \in BV$ fixed we have to compute

$$\inf_{u \in BV} L(u, p) = \lambda \int_{\Omega} f(x)p(x)dx + \inf_{u \in BV} -\lambda \int_{\Omega} u(x)p(x)dx + |u|_{BV}$$

By (7) and (10) we can estimate

$$K(u) := -\lambda \int_{\Omega} u(x)p(x)dx + |u|_{BV} \geq (-\lambda\|p\|_* + 1)|u|_{BV}.$$

Now if $-\lambda\|p\|_* + 1 \geq 0$ then the infimum of $K(u)$ is 0 since we may choose $u = 0$. On the other hand, for any p with bounded G-norm we always find a u_0 which is a maximum in (6), i.e. it satisfies

$$\int_{\Omega} p(x)u_0(x)dx = \|p\|_*|u_0|_{BV}.$$

For the case $-\lambda\|p\|_* + 1 < 0$ we get with such an u_0

$$-\lambda \int_{\Omega} p(x)u_0(x)dx + |u_0|_{BV} = (-\lambda\|p\|_* + 1)|u_0|_{BV},$$

hence we may choose $u = \alpha u_0$, $\alpha \in \mathbb{R}$ to see that the infimum of $K(u)$ is $-\infty$. This shows

$$\inf_{u \in BV} -\lambda \int_{\Omega} p(x)u(x)dx + |u|_{BV} = \begin{cases} 0 & \text{if } \|p\|_* \leq \frac{1}{\lambda} \\ -\infty & \text{if } \|p\|_* > \frac{1}{\lambda} \end{cases}$$

and (17). Since there always exists a p with $\|p\|_* \leq \frac{1}{\lambda}$ (e.g. $p = 0$), the dual problem can be restricted to the constraint problem (18).

Now we show that the dual problem has a solution. Consider the set

$$Y = \{p \in BV \mid |p|_{BV} \leq 1, \|p\|_* \leq \frac{1}{\lambda}\}.$$

This is not empty since $0 \in Y$. Let p_n be a maximizing sequence in Y for (18). From $|p_n|_{BV} \leq 1$, $\|p_n\|_* < \infty$ and the continuous embedding of BV into L^2 ([1]) it follows that $\|p_n\|_{L^2} \leq C$, with some embedding constant C . Note that

by definition of the G-norm the normalization condition $\int p_n(x)dx = 0$ has to hold if Ω is bounded, and hence we can apply the Poincaré-inequality ([1]) to get the L^2 -bound. If $\Omega = \mathbb{R}^n$, this follows directly from the embedding.

Since p_n is bounded in L^2 it has a weakly convergent subsequence i.e. some $p \in L^2$ exists such that

$$\int_{\Omega} g(x)\tilde{p}_n(x)dx \rightarrow \int_{\Omega} g(x)p(x)dx \quad \forall g \in L^2.$$

Moreover by the weak lower semicontinuity $|p|_{BV} \leq 1$. From (10) we get the estimate $\|p\|_* \leq C\|p\|_{L^2} < \infty$ and by definition in (6) there exists a $g \in BV$ with

$$\int_{\Omega} p(x)g(x)dx = |g|_{BV}\|p\|_*.$$

By weak convergence it follows that

$$\int_{\Omega} p(x)g(x) = \lim_{n \rightarrow \infty} \int_{\Omega} \tilde{p}_n(x)g(x)dx \leq \limsup_{n \rightarrow \infty} |g|_{BV}\|\tilde{p}_n\|_* \leq \frac{|g|_{BV}}{\lambda}.$$

This shows that $\|p\|_* \leq \frac{1}{\lambda}$, thus $p \in Y$. Again by weak convergence we conclude $\lim_{n \rightarrow \infty} \int_{\Omega} f(x)p_n(x) = \int_{\Omega} f(x)p(x)dx$. Since p_n was a maximizing sequence p is a maximizing element. \square

We have shown, that the dual problem has a solution. We now deal with the question how this problem is related to the original one. It is well known in convex analysis [17] that the dual problem is always smaller then the primal: for all $|q|_{BV} \leq 1$, $\|q\|_* \leq \frac{1}{\lambda}$, and $v \in BV$ we have

$$G(q) \leq \sup_{|p|_{BV} \leq 1, \|p\|_* \leq \frac{1}{\lambda}} G(p) \leq \inf_{u \in BV} \sup_{|p|_{BV} \leq 1} L(u, p) = \inf_{u \in BV} M(u) \leq M(v).$$

If the inequality between the dual problem and the primal one is strict, then there is a duality gap. For the Meyer-decomposition this does not happen as we will show in Theorem 3.3:

$$\sup_{|p|_{BV} \leq 1, \|p\|_* \leq \frac{1}{\lambda}} G(p) = \sup_{|p|_{BV} \leq 1} \inf_{u \in BV} L(u, p) = \inf_{u \in BV} \sup_{|p|_{BV} \leq 1} L(u, p) = \inf_{u \in BV} M(u) \quad (19)$$

If (19) holds we are faced with a saddle point problem. A pair (\bar{u}, \bar{p}) which satisfies

$$L(\bar{u}, \bar{p}) = \sup_{|p|_{BV} \leq 1} \inf_{u \in BV} L(u, p) = \inf_{u \in BV} \sup_{|p|_{BV} \leq 1} L(u, p) \quad (20)$$

is a saddle point. In the next lemma we show the existence of a saddle point. Moreover, the u -part of a saddle point is a minimizer of $M(u)$, and vice versa:

Theorem 3.3. *Let $f \in L^2$, $\lambda > 0$, then (19) holds and a saddle point (\bar{u}, \bar{p}) with (20) exists. For any saddle point (\bar{u}, \bar{p}) , \bar{u} is a minimizer of $M(u)$, and for any minimizer \bar{u} of $M(u)$ a \bar{p} exists such that (\bar{u}, \bar{p}) is a saddle point.*

Proof. If (\bar{u}, \bar{p}) is a saddle point, then by definition \bar{u} is a minimizer of $M(u)$. On the other hand, if a saddle point exists, then

$$\max_{|p|_{BV} \leq 1} G(p) = \min_{u \in BV} M(u),$$

and any minimizer \bar{u} of $M(\bar{u}) = \max_{|p|_{BV} \leq 1} G(p)$. For such a \bar{u} we find a \bar{p} such that

$$M(\bar{u}) = L(\bar{u}, \bar{p})$$

hence (\bar{u}, \bar{p}) is a saddle point.

We have to show the existence of a saddle point. We use the theorem in [7, Prop. 2.1, 2.2] (see also Remark 2.1 there). Let us define $L(u, p)$ as in (15) on the space $L^2(\Omega)$: $L(u, p) : L_0^2(\Omega) \times L_0^2(\Omega) \rightarrow \mathbb{R} \cup \{\infty\}$, where $L_0^2(\Omega) = L^2(\Omega)$ if $\Omega = \mathbb{R}^n$ and $L_0^2(\Omega) = \{p \in L^2(\Omega) \mid \int_{\Omega} p(x) dx = 0\}$ for the case Ω being bounded. (According to the remark after (11) we can assume the normalization condition $\int_{\Omega} u(x) dx = 0$). The saddle point problem (20) can then be rewritten

$$\sup_{p \in \mathcal{B}} \inf_{u \in \mathcal{A}} L(u, p) = \inf_{u \in \mathcal{A}} \sup_{p \in \mathcal{B}} L(u, p),$$

with $\mathcal{B} = \{p \in BV \mid |p|_{BV} \leq 1\}$ and $\mathcal{A} = BV$. Both \mathcal{A}, \mathcal{B} are closed, convex and nonempty sets. Moreover $L : \mathcal{A} \times \mathcal{B}$ is obviously convex in u and concave in p (in fact, it is affine in p). We additionally need that $L(u, p)$ is upper semicontinuous in p and lower semicontinuous in u . This latter simply follows from the lower semicontinuity of $|u|_{BV}$ and the continuous embedding of $BV \rightarrow L^2$ [1]. Since $L(u, p)$ is affine in p it is continuous and in particular upper semicontinuous. Moreover \mathcal{B} is bounded in L^2 by the Poincaré-inequality [1]. According to [7] have to show the following coercivity condition:

$$\exists_{p_0 \in \mathcal{B}} : \lim_{u \in \mathcal{A}, \|u\|_{L^2} \rightarrow \infty} L(u, p_0) = \infty. \quad (21)$$

But we can simply take $p_0 = 0$, with $L(u, p_0) = |u|_{BV}$. Again by the Poincaré-inequality and the normalization condition we have $|u|_{BV} \geq C \|u\|_{L^2}$ for some constant C . This implies the coercivity (21) and using the results in [7] also the existence of a saddle point. \square

The previous theorem also proves the existence of a minimizer for the functional (11). Moreover the formulation as a saddle point problem is very useful, as it automatically give a lower bound $G(p)$ for the functional. In the next theorem we state the optimality conditions for a saddle point in terms of extremal pairs, similar to the optimality conditions for the ROF-model in Theorem 2.2:

Theorem 3.4. *(\bar{u}, \bar{p}) is a saddle point for $L(u, p)$ if and only if it satisfies the*

following conditions

$$|\bar{u}|_{BV} = \lambda \int_{\Omega} \bar{u}(x) \bar{p}(x) dx \quad (22)$$

$$\|f - \bar{u}\|_* = \int_{\Omega} (f(x) - \bar{u}(x)) \bar{p}(x) dx \quad (23)$$

$$\|\bar{p}\|_* \leq \frac{1}{\lambda} \quad (24)$$

$$|\bar{p}|_{BV} \leq 1. \quad (25)$$

Proof. According to the definition, for a saddle point (\bar{u}, \bar{p}) , $|\bar{p}|_{BV} < 1$, $\|\bar{p}\|_* \leq \frac{1}{\lambda}$ has to hold together with

$$G(\bar{p}) = L(\bar{u}, \bar{p}) = M(\bar{u}).$$

The second equation $L(u, p) = M(u)$ is equivalent to

$$\bar{p} = \operatorname{argmax}_{|p|_{BV} \leq 1} \int_{\Omega} (f(x) - \bar{u}(x)) p(x) dx,$$

which means that $(f - \bar{u}, \bar{p})$ are an extremal pair. By (8) this is equivalent to (23) and (25). The first equation $G(\bar{p}) = L(\bar{u}, \bar{p})$ is equivalent to (22) and (24). \square

Theorem 3.4 can be rephrased in terms of extremal pairs:

Corollary 3.5. *\bar{u} is a minimizer of $M(u)$ if a \bar{p} exists such that both $(\bar{p}, f - \bar{u})$ and (\bar{u}, \bar{p}) are extremal pairs.*

This shows that the optimality conditions are very similar to the ROF-case. For the latter $(\bar{u}, f - \bar{u})$ has to be an extremal pair. For the Meyer-decomposition, \bar{u} , and the texture $f - \bar{u}$ are not directly coupled, but by an additional function \bar{p} .

Theorem 3.4 allows us to conclude certain properties of the optimal decomposition. Some of these were already investigated in [11]. We will deal with the question when the decomposition $f = u + v$ is trivial:

Definition 3.6. *We say $f = u + v$ is a trivial decomposition if either $u = 0$ or $v = 0$.*

The following theorem gives an answer, when a Meyer-decomposition is trivial.

Theorem 3.7. *Let $f \in L^2$, $f \neq 0$, $\lambda > 0$, and let \bar{p} be a maximizer of the functional $G(p)$, i.e.*

$$\bar{p} = \operatorname{argmax}_{\|p\|_* \leq \frac{1}{\lambda}, |p|_{BV} \leq 1} \lambda \int_{\Omega} p(x) f(x) dx.$$

If $\|\bar{p}\|_ < \frac{1}{\lambda}$ then the Meyer-decomposition is trivial with $u = 0$. If $|\bar{p}|_{BV} < 1$ then the Meyer-decomposition is trivial with $v = 0$.*

Proof. The optimality condition (22) can only hold if $\|\bar{p}\|_* = \frac{1}{\lambda}$ or $\bar{u} = 0$, this proves the first assertion. If $|\bar{p}|_{BV} < 1$ and $f - \bar{u} \neq 0$ we define $q := \frac{\bar{p}}{|\bar{p}|_{BV}}$, which satisfies $|q|_{BV} = 1$. Form (23) we can conclude that

$$\int_{\Omega} (f(x) - \bar{u}(x))q(x)dx > \|f - \bar{u}\|_*,$$

which contradicts the definition of the G-norm (6). Hence the texture part has to vanish: $f - \bar{u} = 0$. \square

This shows that the only interesting case occurs when both $|p|_{BV} = 1$ and $\|p\|_* = \frac{1}{\lambda}$ are satisfied for the dual problem.

Theorem 3.7 is an extension of some results proven in [11]. There the authors show that for $\Omega = \mathbb{R}^n$ and $\lambda < 4\pi$ the Meyer-decomposition is trivial with $u = 0$. This can be easily deduced from Theorem 3.7 and the embedding (10). In fact, $|p|_{BV} \leq 1$ implies that $\|p\|_* \leq C$, where C is the embedding constant of $BV \rightarrow G$. For $\Omega = \mathbb{R}^n$ this constant is $\frac{1}{4\pi}$. Hence, if $\frac{1}{\lambda} > C$, then $|p|_* = \frac{1}{\lambda}$ can never hold and the decomposition is trivial. In [11] the author also find a criteria when a trivial decomposition with $v = 0$ occurs. This is the case for so-called plain images and λ sufficiently large. Let us give the definition of a plain image according to [11]:

Definition 3.8. *A function $f \in BV$ is a plain image if a function g exists such that*

$$\int_{\Omega} f(x)g(x)dx = |f|_{BV} \text{ and } \|g\|_* = 1 \quad (26)$$

The notion of plain images is useful, because for them the BV -norm can be expressed via the $\|\cdot\|_*$ -norm. For a plain image f we have from (7) and (26)

$$|f|_{BV} = \sup_{\|g\|_* \leq 1} \int_{\Omega} f(x)g(x)dx.$$

Note that this equation is not true for an arbitrary f since BV is not a reflexive Banach space.

The observation of Haddard and Meyer is that for plain images f with g as in (26) and $\lambda > |g|_{BV}$ the decomposition is trivial with $v = 0$. Again this can be easily proven by Theorem 3.7. Taking $p = \frac{g}{|g|_{BV}}$ gives an element with $|p|_{BV} = 1$ and $\|p\|_* = \frac{1}{|g|_{BV}} \leq \frac{1}{\lambda}$, but by (26) we get

$$G(p) = \lambda \int_{\Omega} f(x)p(x)dx = \lambda \frac{|f|_{BV}}{|g|_{BV}} \geq |f|_{BV} = M(f).$$

Since $G(p)$ is always smaller than $M(u)$ this means that $M(f)$ is the minimum value of (11), yielding $u = f$, $v = 0$.

Finally, we state the optimality conditions in Euler-Lagrange form.

Corollary 3.9. *Let u be a minimizer of the functional (11), which is nontrivial (i.e. $u \neq 0$ and $u \neq f$). Then a $p \in BV$, $p \neq 0$ and a $\mu \in \mathbb{R}$ exists with*

$$-\frac{1}{\lambda} \nabla \cdot \frac{\nabla u}{|\nabla u|} = p \quad (27)$$

$$-\mu \nabla \cdot \frac{\nabla p}{|\nabla p|} = f - u \quad (28)$$

$$|p|_{BV} = 1 \quad (29)$$

$$\mu = \|f - u\|_* \quad (30)$$

$$\|p\|_* = \frac{1}{\lambda}. \quad (31)$$

On the other hand, if these conditions hold for (u, p) and a μ then u is a minimizer of (11).

Proof. According to [14] the partial differential equations (27), (28) are equivalent to the statement that (u, p) and $(p, f - u)$ are extremal pairs. The assertion follows from Theorem 3.4. \square

Note that the last two conditions (30), (31) are not really necessary, but are a consequence of (27) and (28). We included them, because for our numerical approximation of (27) and (28) they do not follow from these equations.

The optimality conditions have a very interesting scaling behavior. We can eliminate the parameter λ in (27) and (28) by an appropriate scaling: Define $\tilde{u} = \frac{1}{\mu}u$, $\tilde{p} = \lambda p$, then the equations can be transformed to

$$-\nabla \cdot \frac{\nabla \tilde{u}}{|\nabla \tilde{u}|} = \tilde{p} \quad (32)$$

$$-\nabla \cdot \frac{\nabla \tilde{p}}{|\nabla \tilde{p}|} = \frac{1}{\mu}f - \tilde{u}, \quad (33)$$

where (29) is now replaced by

$$|\tilde{p}|_{BV} = \lambda \quad (34)$$

and (31) by $|p|_* = 1$.

The form of (32) and (33) indicates that the important parameter for the Meyer-decomposition is not λ but μ . In fact, starting from these equations we can consider μ as the tuning parameter and λ as a derived parameter. If μ is selected and a solution to (32) and (33) is found, then λ can be computed by (34). In this way we arrive with an one-parameter model, but with μ instead of λ as a regularization parameter. From the form of the optimality condition it therefore seems more natural to choose μ as tuning parameter instead of λ . This can also be seen by a reformulation of the Meyer-decomposition in terms of a constraint problem:

This constraint problem is very similar to (11) and is defined the following way (compare [23]): Find u such that

$$|u|_{BV} \rightarrow \min \quad \text{subject to } \|f - u\|_* \leq \mu.$$

It is clear, that the minimum is either attained at the boundary of the constraint set $\|f - u\|_* = \mu$ or in its interior, in which case the minimum u is 0 and the decomposition is trivial. If the minimum is attained at the boundary then we can use a Lagrange multiplier and we end up with an unconstrained optimization problem (11), where λ is the Lagrange multiplier. This observation indicates again that λ should be considered a derived parameter instead of the regularization parameter.

Summing up we can formulate a sketch of our algorithm for computing a decomposition:

Algorithm:

1. Given $f \in L^2(\Omega)$, select $\mu > 0$ as regularization parameter
2. Solve (32) and (33) for \tilde{u} and \tilde{p}
3. Compute the solution $u = \mu\tilde{u}$, and the parameter $\lambda = |\tilde{p}|_{BV}$.

The main computation is of course step 2, solving the optimality conditions. In the next section we discuss the numerical aspects of these equations.

For the sake of completeness let us also state the optimality condition in an alternative form: In the equations (32) and (33) we can eliminate p , which gives a nonlinear fourth order equation for u :

$$\nabla \cdot \frac{\nabla \nabla \cdot \frac{\nabla \tilde{u}}{|\nabla \tilde{u}|}}{|\nabla \nabla \cdot \frac{\nabla \tilde{u}}{|\nabla \tilde{u}|}|} = \frac{1}{\mu} f - \tilde{u}.$$

Although this represents an equation only in u it seems to be quite difficult to deal with from a computational point of view. For the numerics we will stick to the (u, p) formulation of the problem.

Let us compare the optimality conditions above with those for the functional (14). In [4] the authors solve the minimization problem by a projection algorithm: u, v is defined by

$$\bar{u} = f - \bar{v} - P_{\frac{1}{2\lambda}}(f - \bar{v}) \tag{35}$$

$$\bar{v} = P_{\mu}(f - \bar{u}), \tag{36}$$

where P_{λ} denotes the projection onto the set $\{z \in L^2 \mid \|z\|_* \leq \lambda\}$. It is well-known [6] that this projection is another way of finding the ROF minimizer:

$$\bar{u} = \operatorname{argmin} \frac{1}{2\lambda} \|f - u\|_{L^2} + |\nabla u| \Leftrightarrow \bar{u} = f - P_{\lambda}(f).$$

We may rewrite (35), (36) in a more convenient form using ROF-functionals. By the previous identity they are equivalent to

$$\begin{aligned} \bar{u} &= \operatorname{argmin}_u \lambda \|f - \bar{v} - u\|_{L^2}^2 + |u|_{BV} \\ \bar{w} &= \operatorname{argmin}_w \frac{1}{2\mu} \|f - \bar{u} - w\|_{L^2}^2 + |\nabla w| \\ \bar{v} &= f - \bar{u} - \bar{w}. \end{aligned}$$

The optimality conditions for these equations are

$$\begin{aligned} -\frac{1}{2\lambda}\nabla\cdot\frac{\nabla u}{|\nabla u|} + u + v &= f \\ u - \mu\nabla\cdot\frac{\nabla w}{|\nabla w|} + w &= f \\ u + v + w &= f \end{aligned}$$

Eliminating v in the first equation and a scaling $\tilde{u} := \frac{1}{\mu}u$, $\tilde{w} = 2\lambda w$ leads to

$$\begin{aligned} -\nabla\cdot\frac{\nabla\tilde{u}}{|\nabla\tilde{u}|} - \tilde{w} &= 0 \\ \tilde{u} - \nabla\cdot\frac{\nabla\tilde{w}}{|\nabla\tilde{w}|} + \frac{1}{2\lambda\mu}w &= \frac{1}{\mu}f. \end{aligned}$$

It was shown in [4] that for $\lambda \rightarrow \infty$ the solution of (14) approximate the solutions of (11). Now this also can immediately be seen the optimality conditions, since for $\lambda \rightarrow \infty$ the Euler-Lagrange equations coincide. The role of the dual variable p is precisely the role of the 'noise' $w = f - u - v$ in the Aujol-Aubert-Blanc-Ferrand-Chambolle method.

4 Numerical Algorithms

In the following we study the nonlinear saddle point from a numerical point of view with the focus on the optimality conditions in Euler-Lagrange form. Let us define the second order differential operator which is central to Euler-Lagrange equations for BV-related functionals:

$$\mathcal{A}(u) := \nabla\cdot\frac{1}{|\nabla u|}\nabla.$$

To avoid the singularity at $|\nabla u| = 0$, in numerical computations this operator is usually approximated by the ϵ -regularization

$$\mathcal{A}_\epsilon(u) := \nabla\cdot\frac{1}{\sqrt{|\nabla u|^2 + \epsilon}}\nabla. \quad (37)$$

Using this approximation the equations (32), (33) can be written as

$$\mathcal{L}_\epsilon(u, p) \begin{pmatrix} p \\ u \end{pmatrix} := \begin{pmatrix} -\mathcal{A}_\epsilon(p) & I \\ -I & -\mathcal{A}_\epsilon(u) \end{pmatrix} \begin{pmatrix} p \\ u \end{pmatrix} = \begin{pmatrix} \frac{1}{\mu}f \\ 0 \end{pmatrix}.$$

Note that the diagonal parts of \mathcal{L}_ϵ are symmetric positive semidefinite operators which makes \mathcal{L}_ϵ in total an antisymmetric matrix. By a different choice of variables we can easily transform the system into a symmetric but indefinite form. Using $q := -p$ as dual variable gives

$$\tilde{\mathcal{L}}_\epsilon(u, q) \begin{pmatrix} q \\ u \end{pmatrix} := \begin{pmatrix} \mathcal{A}_\epsilon(q) & I \\ I & -\mathcal{A}_\epsilon(u) \end{pmatrix} \begin{pmatrix} q \\ u \end{pmatrix} = \begin{pmatrix} \frac{1}{\mu} f \\ 0 \end{pmatrix}. \quad (38)$$

For a simplified case we may look at the eigenvalues of $\tilde{\mathcal{L}}_\epsilon$: Assume that $\mathcal{A}_\epsilon(u) = \mathcal{A}_\epsilon(q)$, then from the equation for the eigenvectors it follows that the eigenvalues of $\tilde{\mathcal{L}}_\epsilon$ are symmetric around the origin, and the eigenvalues are the same as that of $\mathcal{A}_\epsilon(u)$, but one taken with a positive sign and one taken with a negative sign. This shows some of the difficulties for the numerical computations: Although the matrix $-\mathcal{A}_\epsilon(u)$ is positive semidefinite its eigenvalues might be very small, since in general there is no lower bound on $\sqrt{|\nabla u|^2 + \epsilon}^{-1}$. This makes the equation difficult for convergence analysis and also for computation. In particular, the usual analysis for such saddle point problem requires some sort of ellipticity (the LBB-condition [5]) such as $(\mathcal{A}_\epsilon(q)q, q) \geq C\|q\|^2$. However, such a condition cannot hold, since $(\mathcal{A}_\epsilon(q)q, q) \sim |q|_{BV}$. This indicates that a standard convergence analysis cannot be used. We leave a rigorous convergence analysis of our numerical schemes for future work and only examine convergence numerically.

As we have already pointed out, we consider μ the tuning parameter and not λ . After selecting μ the core of the computation is to solve the discretized nonlinear indefinite problem (38). Of course, there are many possibilities to do this, either by a direct or an iterative solver. The matrix in (38) is a sparse one, hence for iterative solvers the matrix $\tilde{\mathcal{L}}$ does not have to be stored, if only matrix-vector multiplications are used. This is a big advantage especially for large images.

Let us look at some algorithms for solving (38). The simplest one is a Richardson iteration: Denote by r the combined vector $r = (u, p)^T$, and $g := (\mu^{-1}f, 0)^T$ then it is defined as

$$r_{n+1} = r_n + \tau \left(\tilde{\mathcal{L}}_h(r_n)r_n - g \right),$$

with $\tau > 0$. It turns out, that for practical purposes this method is much too slow and not very useful, and does not converge when τ is too large.

A similar variant is the Uzawa-algorithm [2]. Here u_{n+1} is computed by solving the second line in equation (38) and then it is used for the computation of the dual variable q_{n+1} .

$$\begin{aligned} -\mathcal{A}_\epsilon(u_{n+1})u_{n+1} &= -q_n \\ q_{n+1} &= q_n + \tau (\mathcal{A}_\epsilon(q_n)q_n + u_{n+1} - \mu^{-1}f). \end{aligned}$$

Still, this requires the solution of a nonlinear problem for u_{n+1} , which is not easy. In this iteration $\tau > 0$ is a stepsize parameter, which has to be sufficiently small. Unfortunately, the usual convergence theory (see e.g. [7]) does not apply for this iteration. The main obstacle is again the lack of strict positive definiteness of \mathcal{A} . In our numerical computations we observed that the Uzawa iteration does not converge, or requires unreasonable small stepsizes. A main obstacle is that the

matrix $-\mathcal{A}_\epsilon(u_{n+1})$ is not strictly positive definite, hence the equation for u_{n+1} is an almost singular problem. If we compare this equation with the corresponding Euler-Lagrange equation for the ROF problem we see that in the latter a matrix of the form $-\mathcal{A}_\epsilon(u_{n+1}) + I$ has to be inverted, which is not singular due to the added identity matrix. Hence, it seems reasonable to reformulate the Uzawa-iteration into a form with a similar matrix. This can be done by adding the first line in (38) to the second.

As a result we obtain the so called the augmented Lagrangian method [9, 12, 16]:

$$\begin{aligned} -\mathcal{A}_\epsilon(u_{n+1})u_{n+1} + u_{n+1} &= -q_n - \mathcal{A}_\epsilon(q_n) + \mu^{-1}f \\ q_{n+1} &= q_n + \tau (\mathcal{A}_\epsilon(q_n)q_n + u_{n+1} - \mu^{-1}f). \end{aligned} \quad (39)$$

This method looks very appealing, since the first equation is equivalent to an ROF-minimization

$$u_{n+1} = \operatorname{argmin}_{u \in BV} \frac{1}{2} \|(-q_n - \mathcal{A}_\epsilon(q_n) + \mu^{-1}f) - u\|_{L^2}^2 + |u|_{BV},$$

for which standard methods can be used. By definition of (39) we need to solve a nonlinear equation in each step. From a computational point of view this is too expensive, instead we propose to use just one step of solving for u_{n+1} . This leads to the following iteration, where we only have to solve a linear system:

$$\begin{aligned} -\mathcal{A}_\epsilon(u_n)u_{n+1} + u_{n+1} &= -q_n - \mathcal{A}_\epsilon(q_n) + \mu^{-1}f \\ q_{n+1} &= q_n + \tau (\mathcal{A}_\epsilon(q_n)q_n + u_{n+1} - \mu^{-1}f). \end{aligned} \quad (40)$$

We can also improve (41) which is similar to a steepest descent method for an ROF-minimization. Such iterations usually have severe restrictions on the stepsize τ . If we use an implicit method instead we can expect larger stepsize for the cost of solving a linear system. Putting the term involving $\mathcal{A}_\epsilon(q_n)q_n$ on the left hand side and replacing q_n by q_{n+1} in (41) we get the following iteration:

$$\begin{aligned} -\mathcal{A}_\epsilon(u_n)u_{n+1} + u_{n+1} &= -q_n - \mathcal{A}_\epsilon(q_n) + \mu^{-1}f \\ -\tau \mathcal{A}_\epsilon(q_n)q_{n+1} + q_{n+1} &= q_n + \tau (u_{n+1} - \mu^{-1}f). \end{aligned} \quad (42)$$

This iteration is similar to two implicit Euler steps for an minimization of (1). In our numerical computations we observed that the additional computation of a linear equation in (42) gives a faster convergence than (41).

Finally, we can try to solve the whole equation (38) by a fixed point iteration. For $r_n := (u_n, p_n)$ we compute the next iterate r_{n+1} by solving the indefinite linear problem

$$\tilde{\mathcal{L}}_\epsilon(r_n)r_{n+1} = (\mu^{-1}f, 0). \quad (43)$$

The main effort in this iteration is the solution of an linear symmetric indefinite system. Below we will study some algorithm for the linear problem.

Let us now compare the performance of the two inexact augmented lagrangian methods (40)-(41) and (42) with the fixed point iteration (43). In

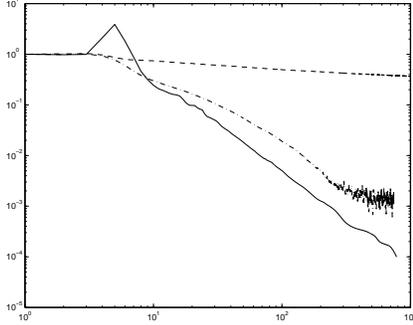


Figure 1: Residual $\mu = 0.1$

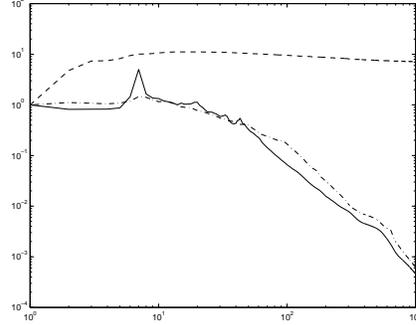


Figure 2: Residual $\mu = 1$

Figure 1 and Figure 2 we show the evolution of the relative residual

$$\text{res}_n := \frac{\|\tilde{\mathcal{L}}_\epsilon(q_n, u_n) - (\mu^{-1}f, 0)\|}{\|\tilde{\mathcal{L}}_\epsilon(q_0, u_0) - (\mu^{-1}f, 0)\|}$$

over the iteration index on a logarithmic scale. For both Figures we choose f the Barbara image with pixel size 128×128 , $\epsilon = 10^{-4}$ in (37) and an initial guess $u_0 = q_0 = 0$. The dashed line represents the iteration (40)-(41) with $\tau = 0.01$, the dashed-dotted line (42) with $\tau = 0.1$ and the full line corresponds to the fixed point iteration (43). Figure 1 shows the result for the parameter choice $\mu = 0.1$ and Figure 2 for $\mu = 1$. It can be seen that the first inexact augmented lagrangian iteration (40)-(41) shows very bad convergence result. However, the second augmented lagrangian variant (42) is almost similar to the fixed point method. However, one fixed-point iteration takes about four times as long as one combined step in (42), which makes this latter method the most efficient one in this case.

Let us now consider the linear indefinite problem for solving one step in the iteration (43). This equation can either be solved directly or by iterative methods. For linear indefinite problems a range of iterative solver have been proposed. The benefit of iterative solver is that the matrix $\tilde{\mathcal{L}}_\epsilon$ does not have to be computed and stored, but only matrix-vector multiplication have to be implemented. On the other hand, storing the matrix does not require much resources, since all the matrices in \mathcal{L}_ϵ are sparse.

The most common iterative procedure for positive definite sparse matrices is the conjugate gradient (CG) method, but since problem (43) is indefinite this method cannot be used in our case. Instead, a couple of generalizations of the CG iteration are available: We consider the Minimal Residual (MINRES), the GMRES method the quasi minimal residual (QMR) and the conjugate gradient method for the normal equation (LSQR). We used these iterations as they are implemented in MATLAB (Version 6.5) and compare the computation time with MATLAB's direct solver for sparse matrices [10]. Our observation is that the iterative solvers cannot compete with MATLAB's direct solver.

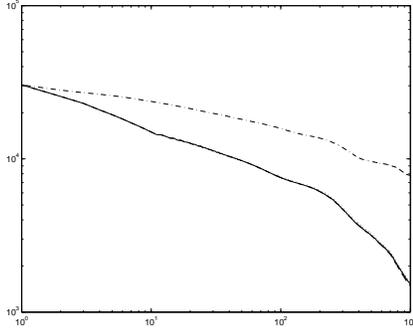


Figure 3: Residual-linear

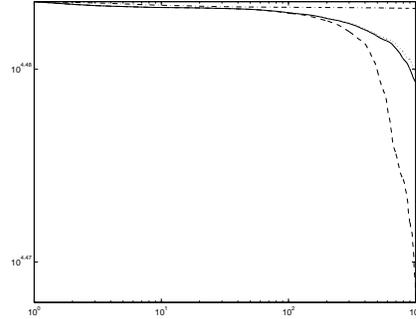


Figure 4: Residual-linear

Method	res ₁₀₀₀ Ex. 1	Time Ex 1	res ₁₀₀₀ Ex. 2	Time Ex 2
direct	$< 10^{-8}$	0.45 s	$< 10^{-8}$	0.41 s
MINRES	0.991	6.7s	0.05	6.9s
GMRES	0.964	472.8s	0.05	425.8s
LSQR	0.999	14.7s	0.25	14.5s
QMR	0.991	13.6 s	0.05	13.5 s

Table 1: Residual and computation time for iterative and direct method

The computations were done with the same picture as for the previous results with the same parameter $\epsilon = 10^{-4}$, $\mu = 0.1$ but on a smaller 64×64 grid. We performed two experiments: for the first one we tried to solve (43) for r_{n+1} , where r_n was set to 0, which makes \mathcal{A}_ϵ the Laplace-operator. The results of the relative residual of (43) versus the iteration number is shown in Figure 3. The full line corresponds to the MINRES-iteration, the dashed one to the GMRES, the dashed-dotted one to the LSQR and the dotted one to the QMR iteration. In Figure 4 we set $r_n = (u_n, q_n)$, where u_n, q_n were the previously computed solution of equation (38). In this case the matrix \mathcal{A}_ϵ corresponds to a second order differential operator with highly discontinuous coefficients. This makes some difficulties for the iterative solvers, as in Figure 4 the convergence is much slower than in Figure 3. Moreover, it can be seen from the results that there is not much difference between the MINRES, the QMR and the GMRES iteration. Only in Figure 4 the GMRES-method shows a slightly faster convergence whereas the LSQR-method is all cases the worst method.

Table 1 shows a comparison of the relative residual after 1000 iterations and the total computation time needed on a Pentium 4/2.2 GHz PC. The second and third row corresponds to the choice $r_n = 0$ as in Figure 3 and the fourth and fifth row to choice of r_n as in Figure 4. It is obvious that the iterative solvers cannot compete with the direct one. Amongst the iterative solver the MINRES is the best one in terms of computation time. Although the GMRES

has a smaller residual for the second example its computation is about 60 times slower than the MINRES method. The GMRES is therefore not suited for this problem, because it is too slow.

We come to the conclusion that iterative solver are not suited for the linear problem. However, we should point out that we do not used any preconditioning, which surely can improve convergence. On the other hand, it is not yet clear how to construct good preconditioners for (43).

Comparing the nonlinear iterations (40)-(43) we prefer to use the augmented Lagrangian (42), over the other methods.

We end this section with some computed decompositions for different parameter μ . In Figure 5 we computed the cartoon part u for the Meyer-decomposition (11) and the ROF-decomposition (1) on the right for different choices of μ . The values were $\mu = 1, 0.5, 0.01, 0.05$ from top to bottom. The regularization parameter λ for the ROF-method was as $\lambda = \frac{1}{2\mu}$. We used the same scaling trick as before also for the the ROF-method, such that the Euler-Lagrange equations for (1) have the form $-\nabla \cdot \frac{\nabla u}{|\nabla u|} + u = \mu^{-1}f$. By Theorem 2.2, [14] and (30) both methods are comparable with this parameter setting in the sense that the residual $f - u$ has the same G-norm in both cases. It can be seen from the pictures that the results quite similar. The Meyer-decomposition seems to give a more precise texture removal as can be seen on parts of the table cloth and the scarf. It also does not suffer from staircasing as the ROF-model, which has problems to recover shaded structures.

5 Conclusion

We have formulated the minimization problem (11) as a saddle point problem, which allows to derive the Euler-Lagrange equations without introducing an additional approximation such as in (12), (13) or (14). Moreover, it leads to a dual problem, which is interesting in its own. Although the dual problem (18) appears to be simpler, we did not make use of it for the numerical computations. The difficulty herein lies in the computation of the norm $\|p\|_*$. Still, it might be interesting to consider a numerical approach based solely on the dual problem. In contrast to alternative approaches we considered a one-parameter model, where we choose μ as the tuning parameter and not λ . In our opinion, this is more natural and it is similar to the regularization parameter for the ROF-method. Finally, we discussed some numerical procedures for solving the optimality conditions. Our favorite method is an inexact augmented Lagrangian method (42), which shows good numerical convergence. Of course, a convergence proof of this scheme is desirable and will be considered in future work. For the linear problem the construction of appropriate preconditioners for iterative method would be necessary to make iterative solvers competitive.

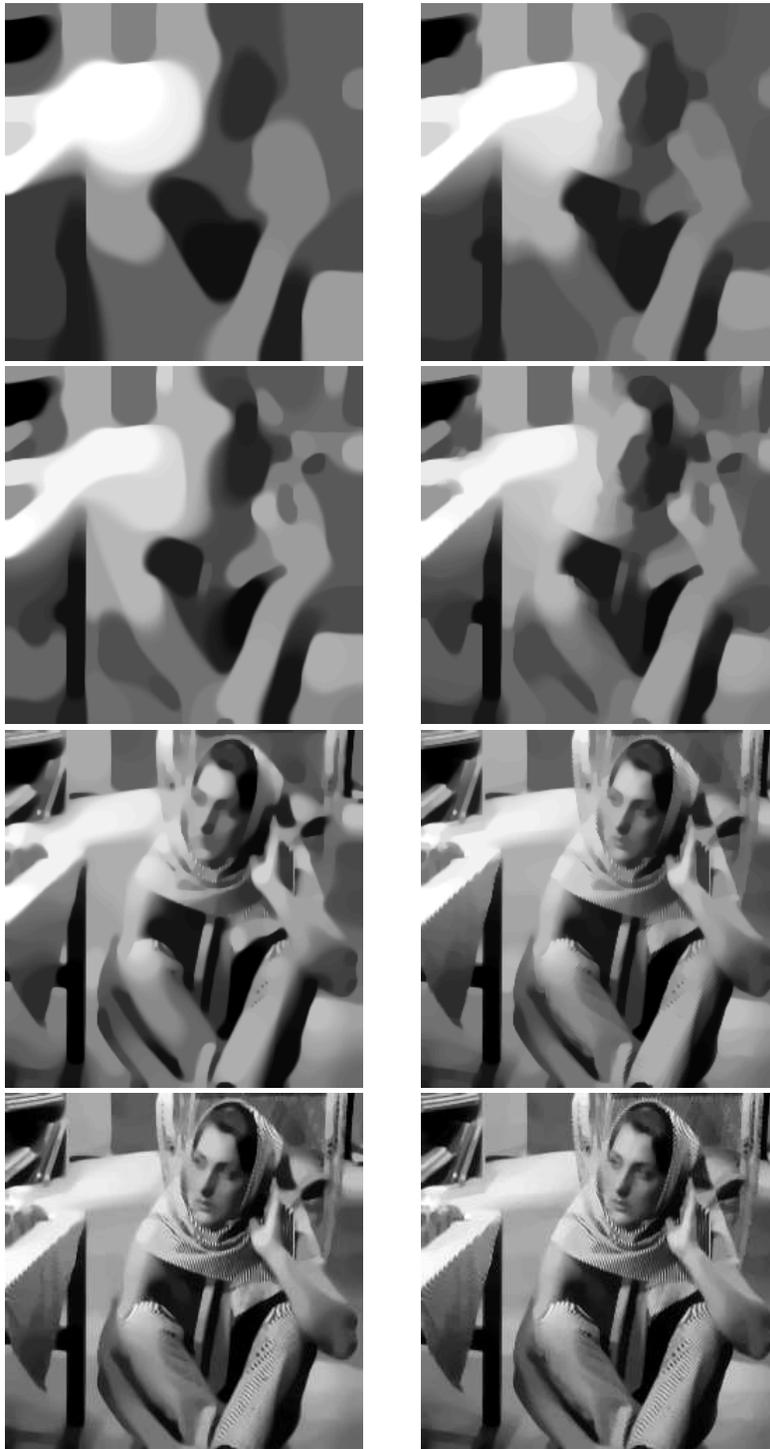


Figure 5: Solution u for Meyer (left) and ROF (right) for $\mu = 1, 0.5, 0.01, 0.05$

References

- [1] L. AMBROSIO, N. FUSCO, AND D. PALLARA, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, New York, 2000.
- [2] K. J. ARROW, L. HURWICZ, AND H. UZAWA, *Studies in linear and nonlinear programming*, Stanford University Press, Stanford, CA, 1958.
- [3] G. AUBERT AND J.-F. AUJOL, *Modeling very oscillating signals. application to image processing*, Appl. Math. Optim, 51 (to appear 2005).
- [4] J.-F. AUJOL, G. AUBERT, L. BLANC-FÉRAUD, AND A. CHAMBOLLE, *Image decomposition into a bounded variation component and an oscillating component*, Journal of Math. Imag. and Vis., 22 (2005), pp. 71–88.
- [5] F. BREZZI AND M. FORTIN, *Mixed and hybrid finite element methods*, Springer-Verlag, New York, 1991.
- [6] A. CHAMBOLLE, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vision, 20 (2004), pp. 89–97.
- [7] I. EKELAND AND R. TÉMAM, *Convex analysis and variational problems*, vol. 28, SIAM, Philadelphia, PA, 1999.
- [8] S. ESEDOĞLU AND S. J. OSHER, *Decomposition of images by the anisotropic Rudin-Osher-Fatemi model*, Comm. Pure Appl. Math., 57 (2004), pp. 1609–1626.
- [9] M. FORTIN AND R. GLOWINSKI, *Augmented Lagrangian methods*, vol. 15 of Studies in Mathematics and its Applications, North-Holland Publishing Co., Amsterdam, 1983.
- [10] J. R. GILBERT, C. MOLER, AND R. SCHREIBER, *Sparse matrices in MATLAB: design and implementation*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 333–356.
- [11] A. HADDARD AND Y. MEYER, *Variational methods in image processing*, preprint, CAM-report, UCLA, 2004.
- [12] M. R. HESTENES, *Multiplier and gradient methods*, J. Optimization Theory Appl., 4 (1969).
- [13] S. KINDERMANN, S. OSHER, AND J. XU, *Denoising by BV-duality*, preprint, CAM-report, UCLA, April, 2005.
- [14] Y. MEYER, *Oscillating patterns in image processing and nonlinear evolution equations*, vol. 22 of University Lecture Series, American Mathematical Society, Providence, RI, 2001.

- [15] S. OSHER, A. SOLÉ, AND L. VESE, *Image decomposition and restoration using total variation minimization and the H^{-1} norm*, Multiscale Model. Simul., 1 (2003), pp. 349–370.
- [16] M. J. D. POWELL, *A method for nonlinear constraints in minimization problems*, in Optimization (Sympos., Univ. Keele, Keele, 1968), Academic Press, London, 1969, pp. 283–298.
- [17] R. T. ROCKAFELLAR, *Convex analysis*, Princeton University Press, Princeton, N.J., 1970.
- [18] L. RUDIN, P.-L. LIONS, AND S. OSHER, *Multiplicative denoising and deblurring: theory and algorithms*, in Geometric level set methods in imaging, vision, and graphics, Springer, New York, 2003, pp. 103–119.
- [19] L. I. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms.*, Physica D, 60 (1992), pp. 259–268.
- [20] G. STRANG, *L^1 and L^∞ approximation of vector fields in the plane*, in Nonlinear partial differential equations in applied science (Tokyo, 1982), vol. 81 of North-Holland Math. Stud., North-Holland, Amsterdam, 1983, pp. 273–288.
- [21] L. VESE AND S. OSHER, *Modeling textures with total variation minimization and oscillating patterns in image processing.*, J. Sci. Comput, 19 (2003), pp. 553–572.
- [22] ———, *Image denoising and decomposition with total variation minimization and oscillatory functions*, J. Math. Imaging Vision, 20 (2004), pp. 7–18.
- [23] W. YIN, D. GOLDFARB, AND S. OSHER, *Total variation based image cartoon-texture decomposition*, preprint, CAM-report, UCLA, April, 2005.