

Graph Cuts for Curvature based Image Denoising

Egil Bae, Juan Shi, Xue-Cheng Tai

Abstract

Minimization of total variation (TV) is a well known method for image denoising. Recently, the relationship between TV minimization problems and binary MRF models has been much explored. This has resulted in some very efficient combinatorial optimization algorithms for the TV minimization problem in the discrete setting via graph cuts. To overcome limitations such as staircasing effects of the relatively simple TV model, variational models based on higher order derivatives have been proposed. The Euler's elastica model is one such higher order model of central importance, which minimizes the curvature of all level lines in the image. Traditional numerical methods for minimizing the energy in such higher order models are complicated and computationally complex. In this work we will present an efficient minimization algorithm based on graph cuts for minimizing the energy in the Euler's elastica model, by simplifying the problem to that of solving a sequence of easy graph representable problems. This sequence has connections to the gradient flow of the energy function, and converges to a minimum point. The numerical experiments show that our new approach is more effective in maintaining smooth visual results while preserving sharp features better than TV models.

Index Terms

higher order models, image denoising, curvature, total variation, binary MRF models, graph cuts.

Egil Bae, Department of Mathematics, University of Bergen, Norway. (Egil.Bae@math.uib.no)

Juan Shi, Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore. (shij0004@e.ntu.edu.sg)

Xue-Cheng Tai, Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore and Department of Mathematics, University of Bergen, Norway. (tai@mi.uib.no)

I. INTRODUCTION

Image denoising is the problem of recovering a true image u from an observed noisy image u^0 . In this work, we assume the original image u has been perturbed by additive gaussian noise η :

$$u^0 = u + \eta.$$

The variational approach is an important paradigm for solving image denoising problems when the image is defined on the continuous domain. Total variation is a powerful notion in such variational problems. The ROF (Rudin-Osher-Fatemi) model was the first image denoising model for 2D-noisy images with additive gaussian white noise using total variation as a regularization term [26]. The main advantage of total variation is the ability to preserve discontinuities. Several numerical methods have been developed for minimizing the energy in such TV-based models, and most of them are based on solving the corresponding Euler-Lagrange equations in the continuous setting. Recently, an equivalence between the anisotropic total variation denoising model and a class of binary second order MRF (markov random field) models has been established in the discrete and quantized setting [8], [9], [4], which has resulted in some very efficient algorithms. In case of L gray values, an optimum can be obtained by solving $\log_2(L) - 1$ binary second order MRFs independently. Such binary MRFs have been studied extensively in computer vision and it is well known that they can be solved very efficiently by graph cuts. As a result, graph cuts based algorithms can be used to efficiently find exact global minima of TV models in the discrete setting.

In spite of the simplicity of TV-based models, they have some disadvantages. Most notably is the so called staircasing effect: piecewise constant images are favored over piecewise smooth images. To counteract these disadvantages, more sophisticated models have recently been developed, where the energy functionals depend on higher order derivatives; see [30], [19], [21], [23], [22], [7], [6], [32], [31], etc. Models that minimize curvature based functionals have been demonstrated to perform particularly well. Of central importance is the Euler's elastica model [23], [22], [7]. In contrast to the total variation model, which minimizes the length of each level line in the image, Euler's elastica minimizes the total curvature of each level line in the image. The Euler's elastica model can also preserve discontinuities.

The main disadvantage of such higher order methods is the difficulty and complexity of

computation. Traditionally the Euler-Lagrange or gradient descent equations are derived. These are nonlinear partial differential equations (PDEs) of order four or even higher. Numerically solving these equations is a time consuming process. Another disadvantage of these methods is the possibility of getting stuck in a local minimum, since the curvature based functionals are not convex. For the denoising application this is not such a big drawback since a good initialization, the noisy input image, is always available.

In this work we will propose an algorithm based on graph cuts for efficiently minimizing the energy in the Euler's elastica model. In the discrete setting, this model has the form of a higher order MRF. In contrast to second order MRFs, these are NP-hard and can generally not be solved exactly. Not much work has been devoted to higher order MRFs so far, however in the last few years some work has appeared much of which was developed simultaneously with ours. Some recent work can be found in [13], [25], [14]. Very recently a technique was also proposed in [17] based on dual decomposition and message passing [18]. This method will in general find an approximate solution by solving a simpler relaxed problem. However, the curvature based models we study do not fit directly into their framework. A linear relaxation approach for region based segmentation models with curvature regularity was proposed in [28]. This approach will also find an approximate solution by solving a simpler linear programming relaxation. This approach could potentially be adapted to denoising by increasing the size of the linear program (L -times as large in case of L gray levels). However, the computational complexity would be quite heavy compared to ours, due to the size of the linear program. The same authors also presented an interesting technique for globally optimizing functions of curvature of a curve in case there are no data term (e.g. for edge based segmentation models in 2D) by minimum ratio cycles [27]. However, this result does not apply for energy functionals with a separate data fitting term.

The curvature based models we study do not directly fit into the framework of [17]. One reason for this is that in the discrete setting, the interaction term in the Euler's elastica model consists of a multiplication of one higher order factor and one second order factor. Curvature is also very complex and nonlinear in the discrete setting. Our approach is different, we instead show that the problem can be simplified to that of solving a sequence of much easier problems, in the form of second order MRFs. The sequence of solutions to these simpler problems has connections to the gradient flow of the energy function and will converge to a local minimum. We claim our approach to be more efficient than previous approaches like [17], [28] due to the

underlying graph cuts minimization technique and logarithmic complexity in the number of gray values.

Like all other curvature based approaches, ours cannot be guaranteed to find a global minimum. The solution may to some extent depend on the initialization. However, for the denoising application, the noisy input image is available as a good initialization. Furthermore, numerical experiments indicate that our algorithm for the Euler's elastica model is rather robust to initialization. Our numerical results also show that the minimization of functionals related to curvature prevents shrinking effects and staircasing effects while keeping sharp edges and smooth interior, compared to the total variation denoising model. Restoration results are generated with higher Peak Signal to Noise Ratio (PSNR) values than the ROF model.

The rest of the paper is organized as follows: In section II we review total variation and graph representable energy functions. We show how graph cuts can be used to efficiently minimize these energy functions. In section III we propose an algorithm for minimizing the Euler's elastica model. Based on our general formulation of energy functional, we can make use of the connection between minimization problems and binary MRFs, and solve the problem via graph cuts. In section IV, we carry out numerical experiments with our method, and compare them to the results of TV in several aspects such as Peak Signal to Noise Ratio (PSNR) values, residual images and visual effects.

II. TOTAL VARIATION AND GRAPH REPRESENTABLE ENERGY FUNCTIONALS

In this section we discuss algorithms for total variation and graph representable second order MRFs, that will be used as substeps in our algorithm for minimizing higher order MRFs. Assume the input image u^0 is defined on a continuous domain Ω . In this work we focus on two dimensional problems where $\mathbf{x} = (x, y) \in \mathbb{R}^2$, but our algorithms and results can easily be generalized to higher dimensions. To recover a denoised image u , can be formulated as a minimization problem with constraints such as $\int_{\Omega} (u(\mathbf{x}) - u^0(\mathbf{x}))^2 d\mathbf{x} \leq \sigma^2 |\Omega|$, where σ^2 denotes a bound on the noise variance and $|\Omega|$ is the area of image domain. The total variation functional of u , defined as $TV(u) = \int_{\Omega} |\nabla u|$, was first used as a regularization term for image denoising in the ROF model taking the form

$$\min_{u \in BV(\Omega)} E^{TV}(u) = a \int_{\Omega} |\nabla u(\mathbf{x})| + \int_{\Omega} (u(\mathbf{x}) - u^0(\mathbf{x}))^2 d\mathbf{x}. \quad (1)$$

The parameter $a > 0$ balances the TV regularization term and the fitting term. The space BV of functions of bounded variation is defined as $BV(\Omega) = \{u \in L^1(\Omega) : TV(u) < \infty\}$. There are two variants of the total variation term, the anisotropic variant:

$$TV_1(u) = \int_{\Omega} |\nabla u|_1 = \int_{\Omega} |u_x| + |u_y| dx,$$

and the isotropic variant:

$$TV_2(u) = \int_{\Omega} |\nabla u|_2 = \int_{\Omega} \sqrt{u_x^2 + u_y^2} dx.$$

The anisotropic variant is graph representable and can be efficiently minimized by graph cuts as explained later. However, $TV_1(u)$ has the disadvantage of penalizing stronger in horizontal and vertical directions, i.e. it is not rotationally invariant. This may lead to some blocky artifacts in the processed image, which are often called metrication errors in the discrete setting.

Both $TV_1(u)$ and $TV_2(u)$ satisfy coarea formulae, which will be used extensively in the following sections. For a positive function u of bounded variation, we define the functions of upper level sets:

$$\theta^\ell = \begin{cases} 1 & \text{if } u \geq \ell, \\ 0 & \text{else.} \end{cases} \quad (2)$$

The coarea formula is given by

$$\int_{\Omega} |\nabla u(\mathbf{x})| d\mathbf{x} = \int_{\ell=0}^L \int_{\Omega} |\nabla \theta^\ell(\mathbf{x})| d\mathbf{x} d\ell, \quad (3)$$

where the image $u(\mathbf{x})$ is a positive function and L is the maximum value of u .

Let us also mention a generalization of the above functional which is graph representable and will be used in later sections

$$E^{TVg}(u) = \int_{\Omega} f^1(\mathbf{x})|u_x| + f^2(\mathbf{x})|u_y| d\mathbf{x} + \int_{\Omega} g(u(\mathbf{x}), \mathbf{x}) d\mathbf{x}, \quad (4)$$

where $f^1, f^2 : \Omega \mapsto \mathbb{R}$ are any positive functions and $g : \Omega \times \mathbb{R} \mapsto \mathbb{R}$ is a convex function in u . In case $f^1 = f^2 = a$ and $g(u(\mathbf{x}), \mathbf{x}) = |u(\mathbf{x}) - u^0(\mathbf{x})|^2$, (4) reduces to the anisotropic TV denoising model. The coarea formula for this general functional is given by

$$\int_{\Omega} f^1(\mathbf{x})|u_x| + f^2(\mathbf{x})|u_y| d\mathbf{x} = \int_{\ell=0}^L \int_{\Omega} f^1(\mathbf{x})|\theta_x^\ell| + f^2(\mathbf{x})|\theta_y^\ell| d\mathbf{x} d\ell, \quad (5)$$

where L is the maximum value of u .

A. Discretization of Energy Functional

We now consider the discrete and quantized versions of (1) and (4). To this end, define the uniform grid

$$\Omega_d = \{(i, j) : i = 1, \dots, N, j = 1, \dots, M\}, \quad (6)$$

and let the discrete functions $u_{i,j} : \Omega_d \mapsto \{0, \dots, L-1\}$ and $u_{i,j}^0 : \Omega \mapsto \{0, \dots, L-1\}$ be defined on this grid. $\{0, \dots, L-1\}$ is the set of gray values. In Fig. 1, Ω_d is depicted as the set of filled circles.

The anisotropic total variation model can be written as

$$E_d^{TV}(u) = \sum_{(i,j) \in \Omega_d} |u_{i,j} - u_{i,j}^0|^2 + a \sum_{(i,j) \in \Omega_d} |u_{i+1,j} - u_{i,j}| + |u_{i,j+1} - u_{i,j}|. \quad (7)$$

Let us now focus on the generalized model (4). To express the discrete version of this model it will be useful to define two new sets of grid points

$$\Omega_d^x = \{(i + \frac{1}{2}, j) : i = 1, \dots, N, j = 1, \dots, M\},$$

and

$$\Omega_d^y = \{(i, j + \frac{1}{2}) : i = 1, \dots, N, j = 1, \dots, M\}.$$

In Fig. 1, Ω_d^x is depicted as the set of squares and Ω_d^y is depicted as the set of empty circles. In the discrete setting, the functions $f^1 : \Omega_d^x \mapsto \mathbb{R}$ and $f^2 : \Omega_d^y \mapsto \mathbb{R}$ are defined on Ω_d^x and Ω_d^y respectively and the function $g : \Omega_d \mapsto \mathbb{R}$ is defined on Ω_d . The discrete version of (4) can now be written

$$E_d^{TVg}(u) = \sum_{i,j} g_{i,j}(u_{i,j}) + \sum_{i,j} f_{i+\frac{1}{2},j}^1 |u_{i+1,j} - u_{i,j}| + f_{i,j+\frac{1}{2}}^2 |u_{i,j+1} - u_{i,j}|. \quad (8)$$

Both the models (7) and (8) are anisotropic and use a 4-neighborhood system for the regularization term. It is possible to create "more" isotropic versions by using a larger neighborhood systems, and hence reduce metrication errors in the solutions. However, for simplicity we stick to 4-neighborhood systems in this work.

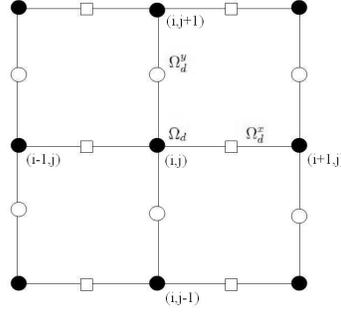


Fig. 1. Grid Definition

B. Equivalent Representation as Set of Binary MRFs

In this section we show how (7) and (8) can be minimized by graph cuts using an algorithm which has a logarithmic complexity in the number of gray values L . In order to simplify notation, we construct an equivalent representation of (8). For each $p = (i, j) \in \Omega_d$, we define the neighborhood system

$$\mathcal{N}_p^4 = \{(i \pm 1, j), (i, j \pm 1)\} \cap \Omega_d.$$

For $p \in \Omega_d$ and $q \in \mathcal{N}_p^4$ the discrete total variation model (7) can be written more compactly using the 4 neighborhood system as

$$E_d^{TV}(u) = \sum_{p \in \Omega_d} |u_p - u_p^0|^2 + \sum_{p \in \Omega_d} \sum_{q \in \mathcal{N}_p^4} \frac{1}{2} w_{pq} |u_p - u_q|. \quad (9)$$

The weights w_{pq} are simply set to a for all neighbors p and q . Using larger neighborhood can generate results that are more isotropic.

For the more general form (8), the notations p^1, p^2 and q^1, q^2 represent the first and second component of p and q respectively. Then (8) can be rewritten as:

$$E_d^{TVg}(u) = \sum_{p \in \Omega_d} g(u_p) + \sum_{p \in \Omega_d} \sum_{q \in \mathcal{N}_p^k} w_{pq} |u_p - u_q|, \quad (10)$$

with the weights w_{pq} given as follows:

$$w_{pq} = \begin{cases} f_{\frac{p^1+q^1}{2}, \frac{p^2+q^2}{2}}^1 & \text{if } p^2 = q^2, \\ f_{\frac{p^1+q^1}{2}, \frac{p^2+q^2}{2}}^2 & \text{if } p^1 = q^1. \end{cases}$$

Darbon et. al. [8], [9] and Chambolle [4] showed independently that the discrete TV model (7) and (9) could be very efficiently minimized via graph cuts, provided the input image is quantized. The same result can also be adapted to the generalized model (8) and (10). In the quantized and discretized case it is assumed that $u : \Omega_d \mapsto [0, 1, \dots, L - 1]$. In this case there are $L - 1$ distinct functions of upper level sets $\ell = 0, \dots, L - 1$, which are point-wise given by

$$\theta_p^\ell = \begin{cases} 1 & \text{if } u_p \geq \ell, \\ 0 & \text{else.} \end{cases} \quad (11)$$

In the discrete setting the coarea formula can be written

$$\sum_{p \in \Omega_d} \sum_{q \in \mathcal{N}_p^k} w_{pq} |u_p - u_q| = \sum_{\ell=0}^{L-1} \sum_{p \in \Omega_d} \sum_{q \in \mathcal{N}_p^k} w_{pq} |\theta_p^\ell - \theta_q^\ell|, \quad (12)$$

where the weights w_{pq} can be derived e.g. by the Cauchy Crofton formula [1]. In this way, the cut based length on the grid graph is close to the Euclidean length of curve, i.e.,

$$\int_{\Omega} |\nabla \theta^\ell(x)| dx d\ell = \sum_{p \in \Omega_d} \sum_{q \in \mathcal{N}_p^k} w_{pq} |\theta_p^\ell - \theta_q^\ell|.$$

If u is a function, it must be single valued at each point $p \in \Omega_d$. In that case the family $\{\theta^\ell\}_{\ell=0}^{L-2}$ is monotonically decreasing: $\theta^{\ell+1} \leq \theta^\ell$, $\forall \ell = 0, \dots, L - 2$. For any such family of monotonically decreasing binary functions, the function u can be recovered by the formula $u_p = \max\{\ell : \theta_p^\ell = 1\}$.

Using these facts it was shown in [4], [8], [9] and [11] that (9) can be written

$$\sum_{\ell=0}^{L-2} E^\ell(\theta^\ell) + \sum_{p \in \Omega_d} |u_p^0|,$$

where

$$\begin{aligned} E^\ell(\theta^\ell) &= \sum_{p \in \Omega_d} \sum_{q \in \mathcal{N}_p^k} w_{pq} |\theta_p^\ell - \theta_q^\ell| + \sum_{p \in \Omega_d} (|\ell + 1 - u_p^0|^2 - |\ell - u_p^0|^2)(1 - \theta_p^\ell) \\ &= \sum_{p \in \Omega_d} \sum_{q \in \mathcal{N}_p^k} w_{pq} |\theta_p^\ell - \theta_q^\ell| + \sum_{p \in \Omega_d} w_p (1 - \theta_p^\ell). \end{aligned} \quad (13)$$

The minimization problem can therefore be written in terms of θ^ℓ as

$$\min_{\{\theta^\ell\}_{\ell=0}^{L-2}} \sum_{\ell=0}^{L-2} E^\ell(\theta^\ell), \quad (14)$$

under the constraints $\theta^{\ell+1} \leq \theta^\ell, \forall \ell = 0, \dots, L-2$. It was shown in [4] that by minimizing each $E^\ell(\theta^\ell)$ independently, these constraints were automatically satisfied. Furthermore, it was observed that each $E^\ell(\theta^\ell)$ has the form of a binary MRF, and could be efficiently minimized via graph cuts as was shown by Greig et. al in 1989 [12]. This leads to several algorithms. Since there is a lot of redundancy between the problems $E^\ell(\theta^\ell)$, it is not necessary to solve the problems sequentially. That would yield an algorithm of linear complexity in the number of gray values. Instead, a dyadic algorithm was proposed in [4], [8], which has a logarithmic complexity in the number of gray values, i.e. at most $\log_2(L)$ binary MRFs had to be solved.

C. Graph Cuts Procedure

Graph cuts has been used in vision problems for a long time. We will briefly review the concept and show how it can be used to efficiently solve binary problems of the form (13). It was first observed in [12] that binary energy functions of the form (10), can be represented on a graph and minimized by max-flow/min-cut algorithms. This construction was also later studied in [2] and [16].

A directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is defined as a set of nodes \mathcal{V} and a set of directed edges \mathcal{E} that connect all the nodes. For 2D grids of the form (6), the set of vertices \mathcal{V} is defined as

$$\mathcal{V} = \{\nu_p \mid p \in \Omega_d\} \cup \{s\} \cup \{t\},$$

where the two distinguished vertices in \mathcal{V} , s and t are called the source and the sink. Then the "cut" $(\mathcal{V}_s, \mathcal{V}_t)$ of the graph is a partition of \mathcal{V} into two sets with $s \in \mathcal{V}_s$ and $t \in \mathcal{V}_t$. The set of edges \mathcal{E} is defined by:

$$\mathcal{E} = \{(\nu_p, \nu_q) \mid p \in \Omega_d, q \in \mathcal{N}_p^k\} \cup \{(s, \nu_p) \mid \forall p \in \Omega_d\} \cup \{(\nu_p, t) \mid \forall p \in \Omega_d\}.$$

$e \in \mathcal{E}$ denotes an edge in \mathcal{E} , and let $c(e)$ denote the "capacity". For a given cut, the set of severed edges \mathcal{C} is defined as:

$$\mathcal{C} = \{(a, b) \in \mathcal{E} \mid a \in \mathcal{V}_s, b \in \mathcal{V}_t\}.$$

The cost of the cut is the energy of this partition as following:

$$E(\mathcal{V}_s, \mathcal{V}_t) = \sum_{e \in \mathcal{C}} c(e).$$

Based on the energy functionals (10) and (13), the costs on the edges from source to every vertex are $c(s, \nu_p) = \max(w_p, 0)$ and the costs from every vertex to the sink are $c(\nu_p, t) = \max(-w_p, 0)$; the costs added on edges between vertices corresponding to neighboring grid points are $c(\nu_p, \nu_q) = w_{\frac{p^1+q^1}{2}, \frac{p^2+q^2}{2}}$. The minimum cut on \mathcal{G} is a minimizer of the energy functional and equals the maximum flow value due to the duality theorem of Ford and Fulkerson [10]. Hence max-flow algorithms can be used to minimize energy functionals of the form (13). In image processing applications such max-flow algorithms have been shown to be extremely efficient [2].

III. MINIMIZATION OF HIGHER ORDER ENERGY FUNCTIONALS INVOLVING CURVATURE

Recently, there have been developed several variational image processing models based on higher order derivatives. Some fourth order PDEs [6], [7], [19], [21], [31] and [32] are especially important. The motivation is to overcome limitations of second order models like total variation. Typically, minimization methods for these models are based on solving their corresponding Euler-Lagrange equations. These equations get increasingly complicated, and numerical solutions are usually rather inefficient. In this work, we make use of graph cuts as a minimization technique. Unfortunately, discretizing the higher order energy functional directly will yield a very difficult combinatorial optimization problem, which is an example of an NP-hard higher order MRF. There has been little work on tackling such higher order MRFs. The recent work of [17] gives details on how to compute approximate solutions to such NP-hard higher order MRFs by using dual decomposition [18]. The idea is to decompose the difficult MRF problem into easier subproblems. Typically each subproblem itself is a higher order mrf over a much smaller graph, even just over one single clique. The subproblems are solved independently. Based on the solutions a master assigns new parameters to each subproblem and they are solved independently over again. This process iterated until convergence.

We show that by instead introducing an artificial time variable, the energy functional can be converted to the graph representable form (8) considered in section II-A. It can therefore be efficiently minimized by graph cuts, as explained in section II-B. Similarly as [17], this approach also simplifies the difficult problem into easier subproblems which are solved iteratively until convergence. We claim our approach is more efficient, mainly because our subproblems can be solved by an algorithm which has a logarithmic complexity in the number of labels. Furthermore,

the underlying graph cut optimization technique has proven to be extremely efficient for practical image processing problems. The complexity of solving the subproblems in [17] is at least linear in the number of labels. Our approach can handle curvature based energy functionals that do not fit into the framework of [17].

The higher order models that minimize some functionals of the curvature of the image [7], [27], [29] and [32] are especially important. Euler's elastica is an important such curvature based model which was first introduced in image processing in [23]. It was studied further in [22], [7], where PDE based optimization methods were developed. In image processing the model can be formulated as the minimization of Euler's elastica of all level curves of the image. The Euler's elastica of a curve Γ is given by the energy

$$E(\Gamma) = \int_{\Gamma} (a + b \cdot |k|^{\beta}(s)) ds, \quad (15)$$

where a and b are two parameters and k is the curvature of Γ at position s . By setting $b = 0$, $E(\Gamma)$ measures the total length of the curve. If $a = 0$, $E(\Gamma)$ measures the total curvature of the curve. Therefore, the Euler's elastica of all level curves of an image u can be written as:

$$\int_{\ell=0}^L \int_{\gamma_{\ell}: u=\ell} (a + b \cdot |k|^{\beta}(s)) ds d\ell. \quad (16)$$

The power β can be set to $\beta = 1$ as in [22], or $\beta = 2$ as in [7]. The choice of $\beta = 1$ allows cracks (sharp corners) of the level curves. By setting $\beta = 2$ differentiation is much easier, such that the Euler-Lagrange equations do not get too complicated. Our approach can handle any choice of β with equal difficulty. For the denoising application we found that $\beta = 1$ tends to give best results. Note that the curvature of the level curve can be expressed as a function of u by

$$k(u) = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right). \quad (17)$$

Using this fact (16) can be expressed more simply as

$$\int_{\Omega} \left(a + b \left| \nabla \cdot \frac{\nabla u}{|\nabla u|} \right|^{\beta} \right) |\nabla u| dx. \quad (18)$$

Euler's elastica can be used as either an image inpainting model or an image denoising model. In the inpainting application, we assume $\tilde{\Omega} \subset \Omega$ is the domain of missing information. The

model is then formulated as

$$\begin{aligned} \min_u \int_{\tilde{\Omega}} \left(a + b \left| \nabla \cdot \frac{\nabla u}{|\nabla u|} \right|^\beta \right) |\nabla u| dx \\ \text{s.t. } u = u_0 \quad \text{on } \partial\tilde{\Omega}, \end{aligned} \quad (19)$$

where $\partial\tilde{\Omega}$ denotes the boundary of $\tilde{\Omega}$.

In this work our emphasis is instead on image denoising, in which case the complete energy functional can be written as

$$\min_u E^{\text{EL}}(u) = \int_{\Omega} |u - u^0|^2 d\mathbf{x} + \int_{\Omega} \left(a + b \left| \nabla \cdot \frac{\nabla u}{|\nabla u|} \right|^\beta \right) |\nabla u| dx. \quad (20)$$

If the parameter b is set to $b = 0$, this model reduces to the usual total variation denoising model, which in view of the discussion above minimizes the total length of all the level lines in the image. If $a = 0$ (20) minimizes the total curvature of all the level lines in the image. Euler's elastica is therefore obviously more sophisticated than total variation, as the resulting level lines will have a more natural appearance. Theoretically, Euler's elastica may also suffer from the staircasing effect, just like total variation, however we found that this model generally gives smoother results. Just as in the TV model, there are anisotropic and isotropic variants of (20) depending on the choice on whether one uses the 1-norm or 2-norm to measure the magnitude of the gradient ∇u . Our algorithm allows to use a mixture with 1-norm in the nominator and 2-norm in the denominator

$$\min_u E^{\text{EL}}(u) = \int_{\Omega} |u - u^0|^2 d\mathbf{x} + \int_{\Omega} \left(a + b \left| \nabla \cdot \frac{\nabla u}{|\nabla u|_2} \right|^\beta \right) |\nabla u|_1 d\mathbf{x}.$$

In this work we will develop graph cuts based methods for minimizing the energy in the models (19) and (20). Neither of these models have the graph representable form considered in Section II. However, by introducing an artificial discrete time variable, we show that a minimum can be obtained by solving a sequence of much easier graph representable problems. Instead of deriving and discretizing the Euler-Lagrange equations or gradient descent equations, we start by discretizing the energy functional itself.

A. Discretization of Curvature based Energy Functionals

The energy functionals (19) and (20) are discretized on the domain Ω_d . We assume further that u and u^0 are quantized, i.e. their function values are constrained to a finite set $u, u^0 : \Omega_d \mapsto$

$[0, 1, \dots, L - 1]$. The discrete gradient ∇u is given by the forward difference scheme:

$$\begin{aligned}\nabla u &= (u^x, u^y), \\ u^x_{i+\frac{1}{2},j} &= u_{i+1,j} - u_{i,j}, \\ u^y_{i,j+\frac{1}{2}} &= u_{i,j+1} - u_{i,j},\end{aligned}\tag{21}$$

where $u^x : \Omega_d^x \mapsto \mathbb{R}$ and $u^y : \Omega_d^y \mapsto \mathbb{R}$. Alternatively, the central difference scheme can be used as a rougher approximation on Ω_d

$$\begin{aligned}\nabla u &= (u^x, u^y), \\ u^x_{i,j} &= (u_{i+1,j} - u_{i-1,j})/2, \\ u^y_{i,j} &= (u_{i,j+1} - u_{i,j-1})/2.\end{aligned}\tag{22}$$

The second order derivatives are discretized as

$$\begin{aligned}u^x_{i,j} &= u_{i+1,j} - 2u_{i,j} + u_{i-1,j}, \\ u^y_{i,j} &= u_{i,j+1} - 2u_{i,j} + u_{i,j-1},\end{aligned}\tag{23}$$

where $u^{xx} : \Omega_d \mapsto \mathbb{R}$ and $u^{yy} : \Omega_d \mapsto \mathbb{R}$.

The curvature k is either discretized with forward difference gradients as

$$k = \left(\frac{u^x}{\sqrt{(u^x)^2 + (u^y)^2}} \right)^x + \left(\frac{u^y}{\sqrt{(u^x)^2 + (u^y)^2}} \right)^y,\tag{24}$$

or with central difference gradients as

$$k = \frac{u^{xx}(u^y)^2 - 2u^x u^y u^{xy} + u^{yy}(u^x)^2}{((u^x)^2 + (u^y)^2)^{\frac{3}{2}}}.\tag{25}$$

Hence $k : \Omega_d \mapsto \mathbb{R}$ is defined on Ω_d . In order to define k on Ω_d^x and Ω_d^y there are several options. One way is to use linear interpolation

$$\begin{aligned}k_{i+\frac{1}{2},j} &= \frac{k_{i+1,j} + k_{i,j}}{2} : \Omega_d^x \mapsto \mathbb{R}, \\ k_{i,j+\frac{1}{2}} &= \frac{k_{i,j+1} + k_{i,j}}{2} : \Omega_d^y \mapsto \mathbb{R},\end{aligned}\tag{26}$$

however, this may result in some blurring around the edges. Another option, which is often used in numerical solutions of PDEs involving curvature is the min-mod [24] discretization

$$\begin{aligned}k_{i+\frac{1}{2},j} &= \text{minmod}(k_{i+1,j}, k_{i,j}) : \Omega_d^x \mapsto \mathbb{R}, \\ k_{i,j+\frac{1}{2}} &= \text{minmod}(k_{i,j+1}, k_{i,j}) : \Omega_d^y \mapsto \mathbb{R}, \\ \text{minmod}(A, B) &= \frac{\text{sign}(A) + \text{sign}(B)}{2} \min(|A|, |B|).\end{aligned}\tag{27}$$

For this discretization to be valid it is required that the time-step in the PDE is set quite low. We therefore experienced this discretization was not very suitable in our algorithm. We instead found the following modification of the min-mod to be more suitable

$$\text{minmod}(A, B) = \begin{cases} |A + B|/2 & \text{if } \text{sign}(A) \neq \text{sign}(B), \\ \frac{\text{sign}(A)+\text{sign}(B)}{2} \min(|A|, |B|) & \text{otherwise.} \end{cases} \quad (28)$$

In our experience, the best results were obtained by central difference gradient and average operator to define k on the middle points.

The discrete version of the Euler's elastica denoising model (20) can then be written with any of the discretizations of k as

$$\min_u E_d^{\text{EL}}(u) = \sum_{(i,j) \in \Omega_d} \left(a + b \cdot |k_{i+\frac{1}{2},j}|^\beta \right) \left| u_{i+\frac{1}{2},j}^x \right| + \left(a + b \cdot |k_{i,j+\frac{1}{2}}|^\beta \right) \left| u_{i,j+\frac{1}{2}}^y \right| + \sum_{(i,j) \in \Omega_d} |u_{i,j} - u_{i,j}^0|^2. \quad (29)$$

Note that the anisotropic gradient factor is discretized with a 4-neighborhood system. A more isotropic version of the elastica model could be derived by using a larger neighborhood system (e.g. 8 neighbors), but for simplicity we stick to 4 neighborhood systems in the work.

We will now define the discrete version of the Euler's elastica inpainting model. Let $\tilde{\Omega}_d \subset \Omega_d$ denote the discrete inpainting domain, the discrete version the inpainting model (19) can then be written

$$\sum_{(i,j) \in \tilde{\Omega}_d} \left(a + b \cdot |k_{i+\frac{1}{2},j}|^\beta \right) \left| u_{i+\frac{1}{2},j}^x \right| + \left(a + b \cdot |k_{i,j+\frac{1}{2}}|^\beta \right) \left| u_{i,j+\frac{1}{2}}^y \right| \quad (30)$$

s.t. $u = u^0$ on $\partial\tilde{\Omega}_d$.

In order to fit the framework of our algorithm, the side constraint must be replaced by a convex data term. Since u is assumed quantized, it is easily seen that (30) can be equivalently rewritten as

$$\min_u E_d^{\text{ELi}}(u) = \sum_{(i,j) \in \tilde{\Omega}_d} \left(a + b \cdot |k_{i+\frac{1}{2},j}|^\beta \right) \left| u_{i+\frac{1}{2},j}^x \right| + \left(a + b \cdot |k_{i,j+\frac{1}{2}}|^\beta \right) \left| u_{i,j+\frac{1}{2}}^y \right| + \lambda \sum_{(i,j) \in \Omega_d \setminus \tilde{\Omega}_d} |u_{i,j} - u_{i,j}^0|^2, \quad (31)$$

provided the parameter λ is sufficiently large compared to a and b .

The discrete models (29) and (31) are now combinatorial optimization problems.

B. Minimization based on Iterative Graph Cuts

It is no doubt the optimization problems (29) and (31) are very complex and difficult to solve exactly, since they do not have the form considered in (8). Furthermore, (29) and (31) do not fit into the framework of [17], because of the nonlinear mixture of pairwise and higher order factors. We will show that the problem can be simplified by introducing an artificial discrete time variable. This is related to the work of [3], [5], where gradient flows of interphases were studied in an integral framework, in case the original energy functional is graph representable in the discrete setting. Our work differs in two main directions: we study images instead of interphases, and we do not require the original energy functional to be graph representable, i.e. we allow more complicated energy functionals involving curvature. Furthermore, we are interested in the final steady state solution, not the evolution sequence itself. We start by the first generalization: let E be a graph representable energy functional of the form (4), depending on images u . The gradient flow of E can be regarded as an infinitesimal sequence of images u^t such that

$$u^{t+dt} = \arg \min_{u: \|u-u^t\| < \varepsilon(dt)} E(u).$$

where $\varepsilon(dt) \rightarrow 0$ as $dt \rightarrow 0$. Choosing the L_2 norm $\|u\|_2^2 = \int_{\Omega} |u|^2 dx$ in the constraint set above will have nice properties, which will soon become apparent. By adding the constraint as a penalty term, this can equivalently be formulated as

$$u^{t+dt} = \arg \min_u E(u) + \gamma(dt) \|u - u^t\|,$$

where $\gamma(dt) \rightarrow \infty$ as $dt \rightarrow 0$, e.g. $\gamma(dt) = \frac{1}{dt}$.

We can now state and prove the following theorem, which is the image counterpart of Theorem 1 given in [3] for interphases.

Theorem 3.1: Given time t_0 , let u^{t_0} be the solution of

$$u^{t_0} = \arg \min_u E(u) + \frac{1}{2(t - t_0)} \|u - u^{t_0}\|_2^2, \quad (32)$$

where $\|u - u^{t_0}\|_2^2 = \int_{\Omega} |u - u^{t_0}|^2 dx$. Then, as $t \rightarrow t_0$

$$u^t = u^{t_0} - \frac{dE}{du}(u^{t_0})(t - t_0) + o(\Delta t), \quad (33)$$

i.e.

$$\frac{\partial u^t}{\partial t} = -\frac{dE}{du}(u^t), \quad (34)$$

where $\frac{dE}{du}$ is the gradient of E with respect to the L_2 inner product.

Proof: Since u^t minimizes (32)

$$0 = \frac{d}{du} \left\{ E(u) + \frac{1}{2(t-t_0)} \|u - u^{t_0}\|^2 \right\} (u^t) = \frac{dE}{du}(u^t) + \frac{1}{(t-t_0)} (u^t - u^{t_0}).$$

By rearranging we get

$$\frac{u^t - u^{t_0}}{t - t_0} = -\frac{dE}{du}(u^t),$$

or, as $t \rightarrow t_0$

$$\frac{\partial u^t}{\partial t} = -\frac{dE}{du}(u^t)$$

■

In the time discrete setting we let Δt denote the step size and $u^n \approx u^{t_0+n\Delta t}$ denote the function u at time step n . Then u^{n+1} is given by

$$u^{n+1} = \arg \min_u E(u) + \frac{1}{2\Delta t} \|u^{n+1} - u^n\|^2 \quad (35)$$

and the discrete gradient flow can then be written

$$u^{n+1} = u^n + \frac{1}{2\Delta t} \frac{\partial E(u^{n+1})}{\partial u},$$

This can therefore be seen as a fully implicit time discretization, where all the factors in $\frac{\partial E}{\partial u}$ are evaluated at the unknown u^{n+1} .

We will now consider problems of the form (29) where the energy functional E itself is not graph representable. By instead making an appropriate semi-implicit time discretization (some factors are handled implicitly and some explicitly), the gradient flow can be approximated by efficiently solving a problem of the form (10) to get from one time step to the next. We also directly move to the spatially discretized models from now on.

We start by considering (18), which is simplest. Letting u^1 be some initial guess and u^0 the input image, we propose the following scheme for minimizing (29): for $n = 1, \dots$, solve

$$\begin{aligned} \min_{u^{n+1}} \sum_{(i,j) \in \Omega_d} \left(a + b \cdot |k_{i+\frac{1}{2},j}^n|^\beta \right) \left| (u_{i+\frac{1}{2},j}^x)^{n+1} \right| + \left(a + b \cdot |k_{i,j+\frac{1}{2}}^n|^\beta \right) \left| (u_{i,j+\frac{1}{2}}^y)^{n+1} \right| \\ + \sum_{(i,j) \in \Omega_d} |u_{i,j}^{n+1} - u_{i,j}^0|^2 + \frac{1}{2\Delta t} \sum_{(i,j) \in \Omega_d} |u_{i,j}^{n+1} - u_{i,j}^n|^2, \end{aligned} \quad (36)$$

The same scheme also applies to the inpainting model (31) by replacing the data term with $\lambda \sum_{(i,j) \in \Omega_d \setminus \tilde{\Omega}_d} |u_{i,j} - u_{i,j}^0|^2$.

Observe that if u^n and k^n at time step n are known, this is an optimization problem of the same form as (8) for the unknown u^{n+1} , with f^1, f^2 and g chosen as

$$\begin{aligned} f_{i+\frac{1}{2},j}^1 &= a + b \cdot |k_{i+\frac{1}{2},j}^n|^\beta : \Omega_d^x \rightarrow \mathbb{R}, \\ f_{i,j+\frac{1}{2}}^2 &= a + b \cdot |k_{i,j+\frac{1}{2}}^n|^\beta : \Omega_d^y \rightarrow \mathbb{R}, \\ g_{i,j} &= |u_{i,j}^{n+1} - u_{i,j}^0|^2 + \frac{1}{2\Delta t} |u_{i,j}^{n+1} - u_{i,j}^n|^2 : \Omega_d \rightarrow \mathbb{R}. \end{aligned} \quad (37)$$

The function $g_{i,j}$ is convex in $u_{i,j}^{n+1}$ as required, being a sum of two convex functions. Therefore, in the spatially discrete setting one can get from one time step to the next by efficiently minimizing $\log_2(L) - 1$ binary MRFs by graph cuts as described in section II-B.

The penalty term of (36) $\frac{1}{2\Delta t} \sum_{i,j} |u_{i,j}^{n+1} - u_{i,j}^n|^2$ is essential for ensuring convergence of the algorithm. Without this term, the algorithm would alternate. As we see, it is incorporated in the algorithm directly in the data term of the second order MRF subproblem. If Δt is low the penalty term is large, therefore the solution u is not allowed make a large displacement from each time step to the next. The relation to the gradient flow holds if the time step Δt is chosen sufficiently small. However, we experienced that the algorithm still converges to the same solutions even if the time step is set quite large. On the other hand, a large choice of Δt may result in oscillations as the sequence u^{n+1} reaches a stationary point. As a compromise, we found it useful to choose a large Δt in the beginning (e.g. first 10-15 iterations) and set Δt lower in the final iterations. It is difficult to find a Δt which completely prevents small oscillations around the stationary point. A useful technique is to add artificial intermediate gray values in the computation, i.e. to compute on a finer label scale. For instance, if an output of 256 gray values is required (i.e. label set: $\{0, 1, \dots, 255\}$), 1024 gray values can be used in the computation (label set in computation: $\{0, 0.25, 0.5, \dots, 254.75, 255\}$). In this case there are 3 artificial gray levels between each true gray value. This strategy aims to force most of the oscillations to occur at the fine scale (adding 3 such artificial labels seems to be sufficient). In the end, the solution is converted to 256 gray values by rounding. This approach will not add much to the cost of the algorithm, since computation with 1024 gray values instead of 256 requires to solve $9 = \log_2(1024) - 1$ binary problems instead of $7 = \log_2(256) - 1$ in each iteration.

To summarize, algorithm 1 iteratively solves a sequence of second order MRF problems until convergence. Each of these MRF problems are converted to $\log_2(L) - 1$ binary MRF problems, where L is the number of gray values, each of which are solved efficiently by graph cuts.

Algorithm 1 Algorithm for Euler's Elastica Model

-
- Step1: Input image u^0 , set $u^1 = u^0$;
 - Step2: For $n = 1, 2, 3\dots$ until converge,
 find u^{n+1} by solving (36) with u^0 and u^n as fixed inputs. This subproblem is a second order MRF with parameters from (37), and is solved via graph cuts by as described in section II-B and II-C;
 - Step3: Output u^{n+1} .
-

Therefore the complexity is logarithmic in the number of gray values. The approach of [17] is also based on iteratively solving a sequence of simpler problems until convergence to an approximate solution of the original problem. However, the complexity of solving the simpler problems increases at least linearly with L as far as we can see.

IV. NUMERICAL EXPERIMENTS

In this section, the algorithm is tested numerically and the results are compared to anisotropic total variation generated by graph cuts. The codes are implemented in C++. For max-flow computation, we have used the implementation of Boykov et. al. reported in [2]. The numerical experiments were performed on an HP laptop with an Intel(R) Core(TM) 2 Duo CPU T7500 @ 2.19GHz, and 2.00GB of RAM. The results are compared visually and by measuring the Peak Signal to Noise Ratio (PSNR) between the ground truth and the output of different methods. The PSNR is defined as:

$$\text{PSNR}(u, v) = 10 \log_{10} \frac{255^2}{\frac{1}{MN} \sum_{i,j} (u_{i,j} - v_{i,j})^2},$$

where $u_{i,j}$ and $v_{i,j}$ denote the pixel values of initial ground truth images and restored images respectively. In addition, the residual image $r = u^0 - v$ where u^0 is the noisy input image and v is the restored output image is shown for some of the experiments.

We start by giving illustrative inpainting examples in Figure 2 which clearly show the advantage of using curvature information over total variation. The inpainting domain $\tilde{\Omega}$ consists of the noisy regions. Random noise has been added in the inpainting domain such that an initialization is available for our method. The same result was also obtained by initializing $u^1 = 255$ (all white) in the inpainting domain. As we see, our algorithm for minimization of Euler's elastica

perfectly connects the level lines such that the missing region is inpainted in the visually most logical manner. Total variation on the other hand, just minimizes the total length of the level lines, which leads to much less reasonable results. The sudden turns in the level lines are not penalized by TV.

Figure 3 also shows the same effect on a more realistic image denoising example. The thin elongated structures are better preserved by the Euler's elastica model. Furthermore, the level lines have a smoother, more natural appearance. Figure 3 (d) and (e) show the set of pixels where the gradient of the image is nonzero, i.e the set of all level lines. The reconstructed image is constant in each white region. As we see, total variation favors solutions that are piecewise constant to a larger degree than Euler's elastica, the so called staircasing effect. In all the examples so far we used central difference gradient in the curvature discretization and the average operation to define the curvature on the middle points (short: central, average in Table 1). The power in the Euler's elastica model is set to $\beta = 1$.

For Lena (512×512) and Pepper (512×512) presented in Figure 5-9, gaussian noise with a variance of $\sigma^2 = 0.01$ was added. In these experiments we have selected the regularization parameters in TV and the curvature based model to optimize results visually. In the Euler's elastica model (Algorithm 1) we set $a = 0$, $b = 90$ and $\beta = 1$, so that the curvature term does all the regularization. The case that $a > 0$ is not so interesting, since that would be a balancing of TV and curvature regularization. Subfigures (d) and (e) show the Euler elastica result with forward discretization of the gradient in the curvature. To define the curvature on the middle points minmod is used in (d) and average is used in (e). In subfigure (f), central difference discretization of the curvature is used together with average operation to define curvature on the middle points. Because of the high resolution, the difference will be more obvious in the electronic version.

In all these experiments we observe that the algorithm for Euler's elastica yields smoother results than total variation. In all the TV results one can clearly notice the staircasing effect: piecewise constant image are favored. The Euler's elastica model may also suffer from the staircasing effect, but to a less degree than TV.

The results are also compared by measuring the Peak Signal to Noise Ratio (PSNR) in Table 1, which shows that the curvature based methods clearly perform best for all the experiments. The residual images for the curvature based methods also include less signal information than total

variation .

We found that Algorithm 1 for Euler’s elastica converged in around 30-40 iterations. This can also be seen from the graphs in Figure 4 where we have plotted the energy $E^{EL}(u^n)$ as a function of the iteration count n for the two images Lena and Pepper. The graphs also show that results generated by Algorithm 1 are quite insensitive to initialization. The red curve depicts energies after initialization with the noisy image $u^1 = u^0$, while the blue curve depicts energies after initialization with the constant image $u^1 = 155$. They both converge to solutions of almost the same energy. For initializations with the noisy image u^0 , most of the change happen during the first 5-10 iterations. The cpu time for our implementation is around 6 seconds per iteration for the lena and pepper image (size 512×512) with 1024 labels, if the max-flow problems are solved from scratch in each iteration. However, since the edge weights only change a little from each iteration to the next one should in practice reuse flow from the previous iteration to get a much faster algorithm. For instance, the work of Kohli et. al. [15] presented such a max-flow algorithm for dynamically changing MRF models. When the changes in the weights are small this algorithm is shown to be significantly faster than traditional max-flow algorithms. However, a careful implementation of such methods is out of the scope of this paper.

Img. No.	Table1: PSNR results comparison ($\sigma^2 = 0.01$)			
	Name	TV	Euler’s Elastica, forward, average	Euler’s Elastica, central, average
Fig. 5	Lena(512×512)	21.86dB	23.46dB	26.59dB
Fig. 8	Pepper(512×512)	24.86dB	26.80dB	28.07dB
Fig. 3	Dragonfly(480×320)	15.71dB	-	18.87dB

V. CONCLUSIONS

In this work we proposed simple and effective algorithms based on graph cuts, which can minimize the energy in the Euler’s elastica model. Based on the connection between TV minimization and binary MRFs, our approach simplifies the minimization problems to that of solving a sequence of simpler problems. The sequence of solutions to these problems converges to a minimum point. Experiments show that shrinking and staircasing effects are prevented in the results generated by our methods. We will apply dynamic max-flow algorithms which reuses

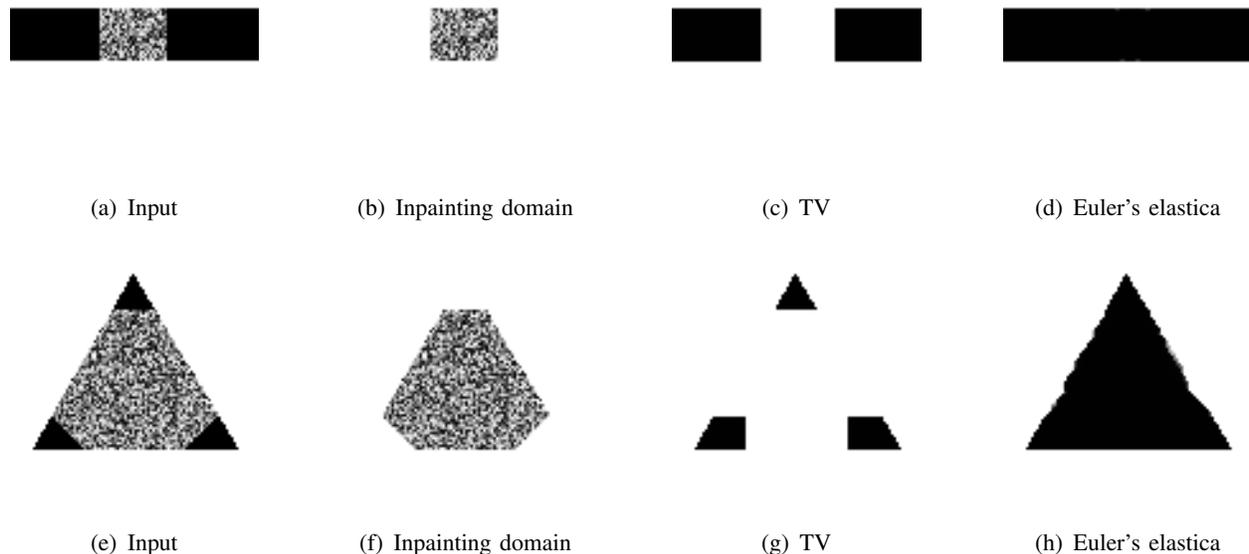


Fig. 2. Inpainting example: The region with random noise denotes the inpainting domain. By minimizing the total curvature of the level lines, our method can fill in the missing regions appropriately. The data term is set to zero in the inpainting domain (the noisy regions). Figure (c) and (g) shows the result of total variation, which minimizes the total length of the level lines. Figure (d) and (h) shows our result of minimizing Euler's elastica with $a = 0$ and using the noisy image as initialization. Some artifacts can be observed due to the simple 4-neighborhood implementation.

flow from each iteration to the next in a future work.

ACKNOWLEDGEMENT

The research has been supported by MOE (Ministry of Education) Tier II project T207N2202 and IDM project NRF2007IDM-IDM002-010 and the Norwegian Research Council (eVita project 166075).

REFERENCES

- [1] Y. Boykov and V. Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. In *Ninth IEEE International Conference on Computer Vision, 2003. Proceedings*, pages 26–33, 2003.
- [2] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1124–1137, 2004.
- [3] Y. Boykov, V. Kolmogorov, D. Cremers, and A. Delong. An integral solution to surface evolution PDEs via geo-cuts. In *9th European Conference on Computer Vision (ECCV). Graz, Austria*, volume 3953, pages 409–422. Springer, 2006.

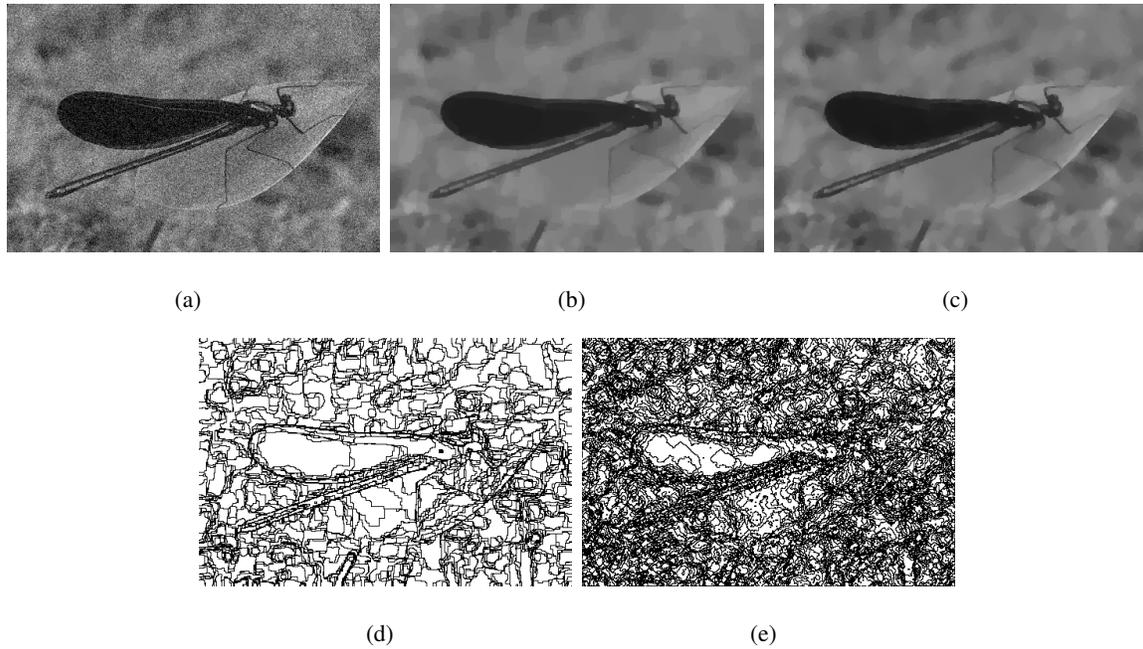


Fig. 3. Denoising results for Dragonfly(480×320) with $\sigma^2 = 0.01$. Thin elongated structures are better preserved by the Euler's elastica model. (a) input noisy image, (b) recovered by TV, (c) recovered by Euler's elastica. Set of pixels with nonzero gradient (level lines): (d) TV, (e) Euler's elastica. The reconstructed image is constant in the white regions, observe that TV favors piecewise constant images to a larger degree than Euler's elastica.

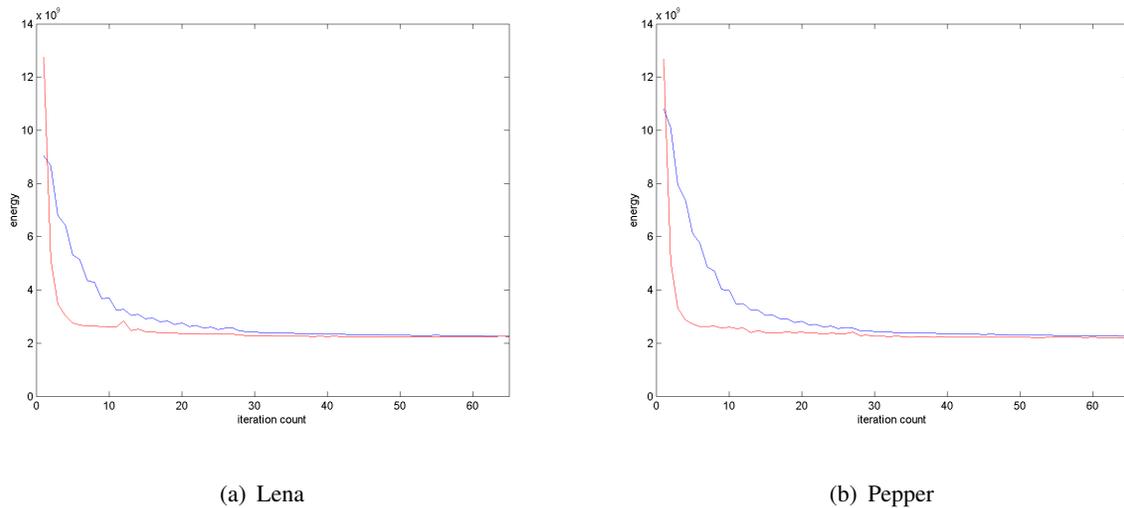


Fig. 4. Energy as a function of the iteration count for Algorithm 1 (Euler's elastica) applied to the Lena image (a) and the Pepper image (b). Red: noisy image $u^1 = u^0$ used as initialization. Blue: constant image $u^1 = 155$ used as initialization. Convergence is reached in around 30-40 iteration for noisy initialization.



(a)

(b)



(c)

(d)



(e)

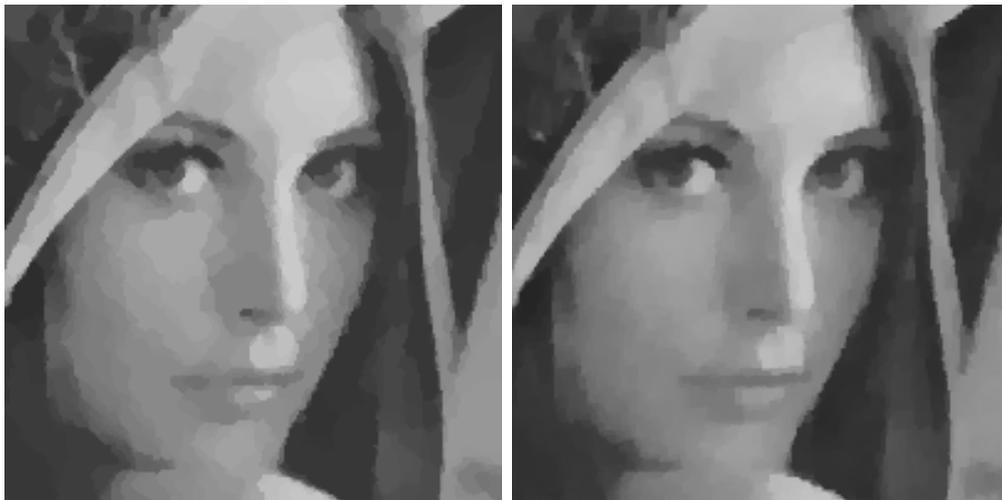
(f)

June 8, 2010
 Fig. 5. Denoising results for Lena(512×512) with $\sigma^2 = 0.01$. (a) original image, (b) input noisy image, (c) recovered image by TV, (d) recovered image by Euler's elastica min-mod forward discretization, (e) recovered image by Euler's elastica average forward discretization, (f) recovered image by Euler's elastica central difference gradient. DRAFT



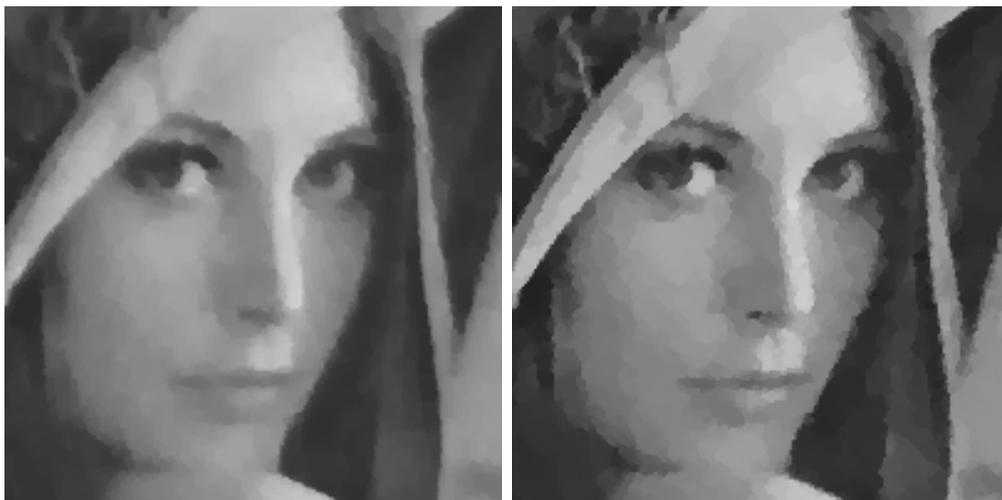
(a)

(b)



(c)

(d)



(e)

(f)

June 8, 2010
 Fig. 6. Denoising results for part of Lena with $\sigma^2 = 0.01$. (a) original image, (b) input noisy image, (c) recovered image by TV, (d) recovered image by Euler's elastica min-mod forward discretization, (e) recovered image by Euler's elastica average forward discretization, (f) recovered image by Euler's elastica central difference gradient. DRAFT

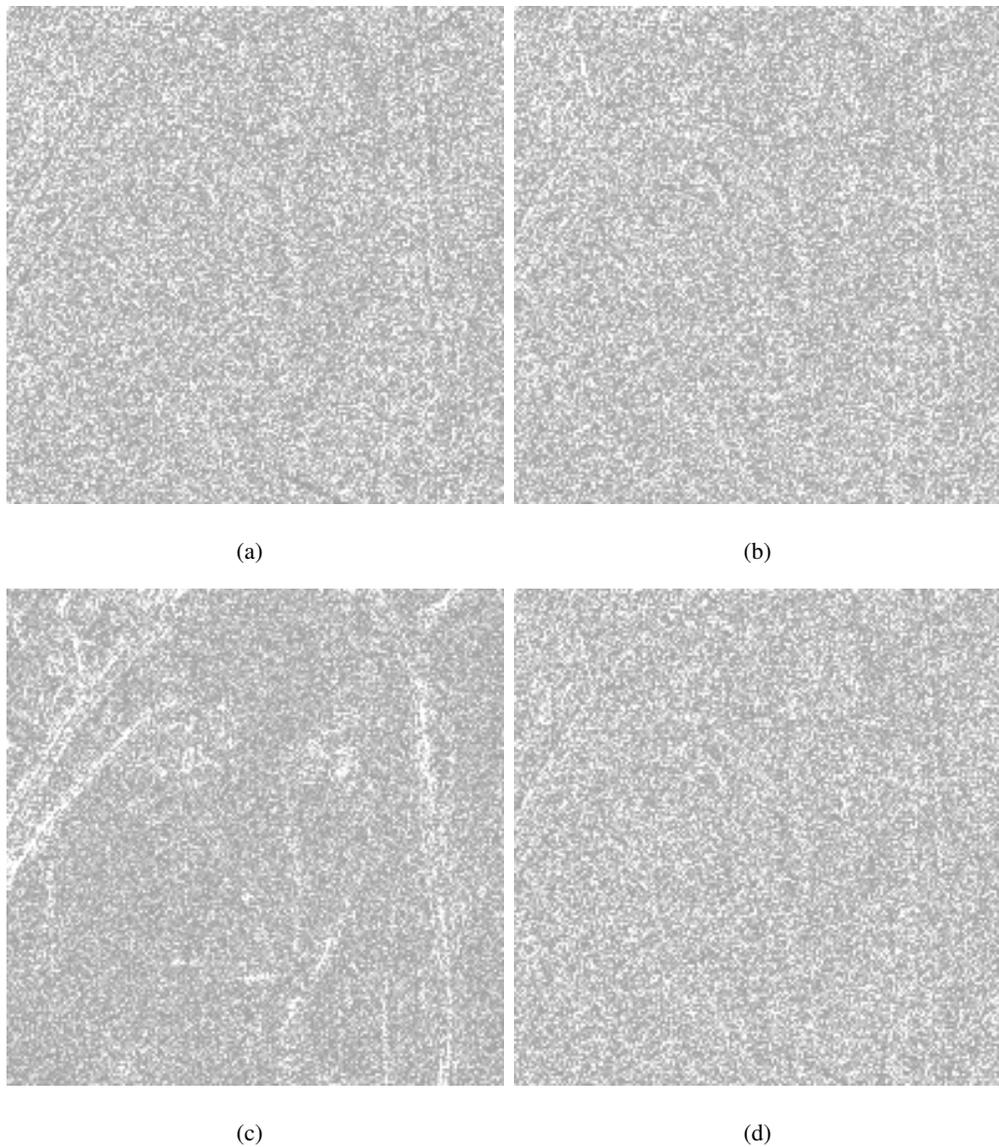
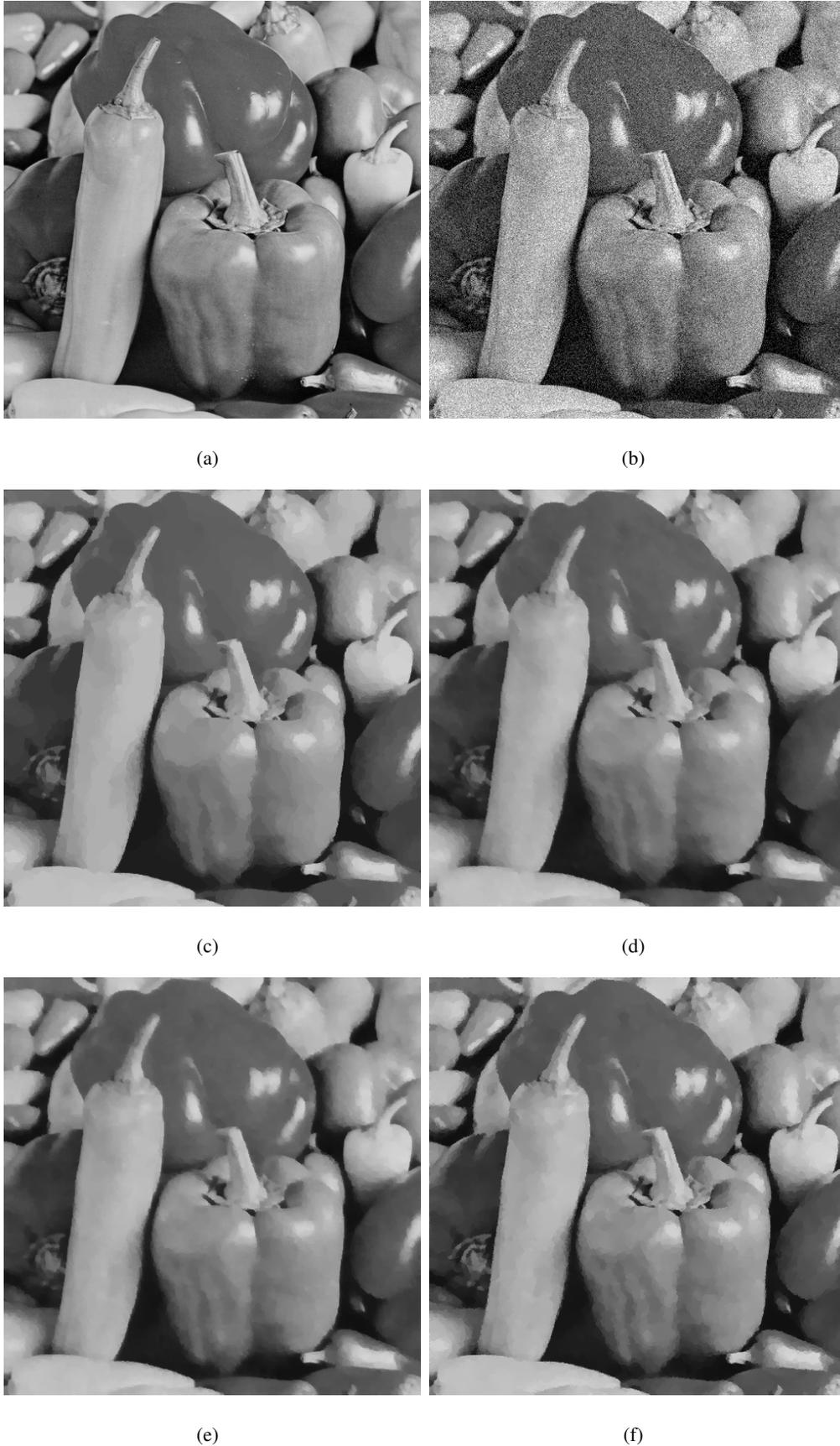
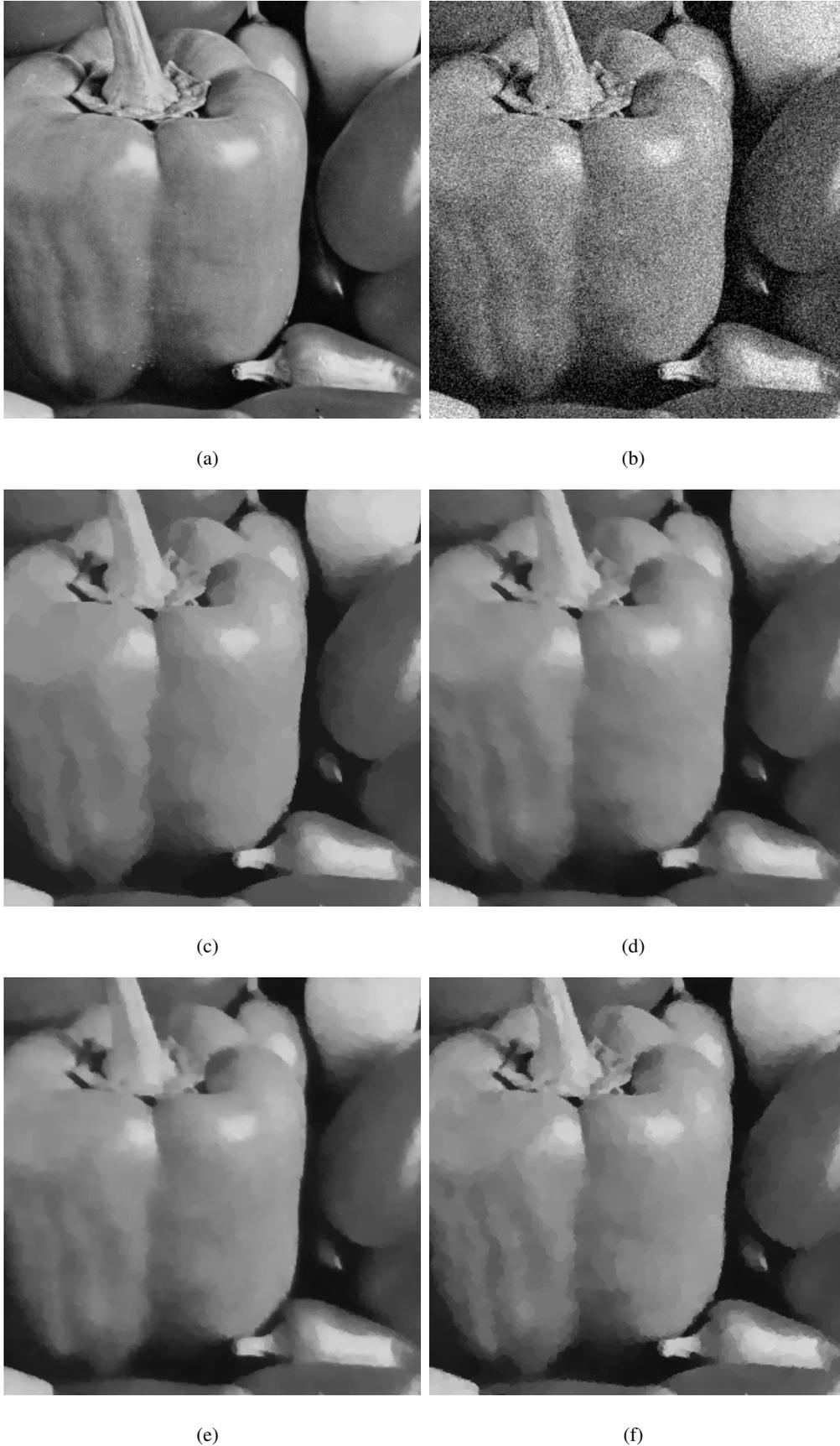


Fig. 7. Difference between reconstructed image and noisy image for part of Lena. (a) TV, (b) Euler's elastica min-mod forward discretization, (c) Euler's elastica average forward discretization, (d) Euler's elastica central difference gradient.

- [4] A. Chambolle. Total variation minimization and a class of binary MRF models. In *Energy Minimization Methods in Computer Vision and Pattern Recognition: 5th International Workshop*, volume 3757, pages 136–152. Springer, 2005.
- [5] A. Chambolle and J. Darbon. On total variation minimization and surface evolution using parametric maximum flows. *International Journal of Computer Vision*, 84(3):288–307, 2009.
- [6] T. Chan, A. Marquina, and P. Mulet. High-order total variation-based image restoration. *SIAM J. SCI. COMPUT.*, 22(2):503–516, 2000.
- [7] T.F. Chan, S.H. Kang, and J. Shen. Euler's elastica and curvature-based inpainting. *SIAM Journal on Applied Mathematics*, pages 564–592, 2002.



June 8, 2010
 Fig. 8. Denoising results for Pepper(512×512) with $\sigma^2 = 0.01$. (a) original image, (b) input noisy image, (c) recovered image by TV, (d) recovered image by Euler's elastica min-mod forward discretization, (e) recovered image by Euler's elastica average forward discretization, (f) recovered image by Euler's elastica central difference gradient. DRAFT



June 8, 2010
 Fig. 9. Denoising results for part of Pepper with $\sigma^2 = 0.01$. a) original image, (b) input noisy image, (c) recovered image by TV, (d) recovered image by Euler's elastica min-mod forward discretization, (e) recovered image by Euler's elastica average forward discretization, (f) recovered image by Euler's elastica central difference gradient. DRAFT



Fig. 10. Set of pixels with nonzero gradient (level lines). (a) Lena reconstructed by TV (from Fig. 5 c), (b) Lena reconstructed by Euler's elastica with central difference gradient (from Fig. 5 f), (c) Pepper reconstructed by TV (from Fig. 8 c), (d) Pepper reconstructed by Euler's elastica with central difference gradient (from Fig. 8 f). The reconstructed image is constant in the white regions, observe that TV favors piecewise constant images to a larger degree than Euler's elastica.

- [8] J. Darbon and M. Sigelle. Image restoration with discrete constrained total variation part I: Fast and exact optimization. *Journal of Mathematical Imaging and Vision*, 26(3):261–276, 2006.
- [9] J. Darbon and M. Sigelle. Image restoration with discrete constrained Total Variation part II: Levelable functions, convex priors and non-convex cases. *Journal of Mathematical Imaging and Vision*, 26(3):277–291, 2006.
- [10] L.R. Ford and D.R. Fulkerson. Flow in networks. *Princeton University Press, Princeton, NJ*, 1962.
- [11] D. Goldfarb and W. Yin. Parametric maximum flow algorithms for fast total variation minimization. *Rice University*

CAAM Technical Report TR07-09, 2007.

- [12] D.M. Greig, B.T. Porteous, and A.H. Seheult. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 271–279, 1989.
- [13] H. Ishikawa. Higher-order clique reduction in binary graph cut. In *CVPR*, pages 2993–3000, 2009.
- [14] Pushmeet Kohli, M. Pawan Kumar, and Philip H. S. Torr. P3 & beyond: Solving energies with higher order cliques. In *CVPR*, pages 1–8, 2007.
- [15] Pushmeet Kohli and Philip H. S. Torr. Dynamic graph cuts for efficient inference in markov random fields. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(12):2079–2088, 2007.
- [16] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):147–159, 2004.
- [17] N. Komodakis and N. Paragios. Beyond Pairwise Energies: Efficient Optimization for Higher-order MRFs. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.
- [18] Nikos Komodakis, Nikos Paragios, and Georgios Tziritas. Mrf optimization via dual decomposition: Message-passing revisited. In *ICCV*, pages 1–8, 2007.
- [19] M. Lysaker, A. Lundervold, and X.C. Tai. Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time. *IEEE Transactions on Image Processing*, 12(12):1579–1590, 2003.
- [20] M. Lysaker, S. Osher, and X.C. Tai. Noise removal using smoothed normals and surface fitting. *IEEE Transactions on Image Processing*, 13(10):1345–1357, 2004.
- [21] M. Lysaker and X.C. Tai. Iterative image restoration combining total variation minimization and a second-order functional. *International Journal of Computer Vision*, 66(1):5–18, 2006.
- [22] Simon Masnou and Jean-Michel Morel. Level lines based disocclusion. In *ICIP (3)*, pages 259–263, 1998.
- [23] D. Mumford. Elastica and computer vision. In *Algebraic Geometry and its Applications*, 1994.
- [24] Stanley Osher and Leonid I. Rudin. Feature-oriented image enhancement using shock filters. *SIAM J. Numer. Anal.*, 27(4):919–940, 1990.
- [25] Carsten Rother, Pushmeet Kohli, Wei Feng, and Jiaya Jia. Minimizing sparse higher order energy functions of discrete variables. In *CVPR*, pages 1382–1389, 2009.
- [26] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1-4):259–268, 1992.
- [27] T. Schoenemann and D. Cremers. Introducing curvature into globally optimal image segmentation: Minimum ratio cycles on product graphs. In *IEEE International Conference on Computer Vision (ICCV). Rio de Janeiro, Brazil*, 2007.
- [28] T. Schoenemann, F. Kahl, and D. Cremers. Curvature regularity for region-based image segmentation and inpainting: A linear programming relaxation. In *IEEE International Conference on Computer Vision (ICCV)*, Kyoto, Japan, 2009.
- [29] T. Tasdizen, R. Whitaker, P. Burchard, and S. Osher. Geometric surface smoothing via anisotropic diffusion of normals. *IEEE Visualization*, pages 125–132, 2002.
- [30] Yu-Li You and Mostafa Kaveh. Fourth-order partial differential equations for noise removal. *IEEE Transactions on Image Processing*, 9(10):1723–1730, 2000.
- [31] Jing. Yuan, Christoph. Schnorr, and Gabriele. Steidl. Total-Variation Based Piecewise Affine Regularization. In *Second International Conference, SSVN. Voss, Norway*, pages 552–564. Springer, 2009.

- [32] Wei. Zhu and Tony. Chan. Image denoising using mean curvature Available in: <http://www.cims.nyu.edu/~wzhu/meancurvature06Nov30.pdf>. 2009.