Global Minimization of the Active Contour Model with TV-Inpainting and Two-Phase Denoising

Shingyu Leung and Stanley Osher

Department of Mathematics, UCLA, Los Angeles, CA 90095, USA
{syleung, sjo}@math.ucla.edu
http://www.math.ucla.edu/~syleung
http://www.math.ucla.edu/~sjo

Abstract. The active contour model [8,9,2] is one of the most well-known variational methods in image segmentation. In a recent paper by Bresson et al. [1], a link between the active contour model and the variational denoising model of Rudin-Osher-Fatemi (ROF) [10] was demonstrated. This relation provides a method to determine the **global** minimizer of the active contour model. In this paper, we propose a variation of this method to determine the global minimizer of the active contour model in the case when there are missing regions in the observed image. The idea is to turn off the L^1-fidelity term in some subdomains, in particular the regions for image inpainting. Minimizing this energy provides a unified way to perform image denoising, segmentation and inpainting.

1 Introduction

Image segmentation, image restoration and image inpainting are a few basic yet important areas in image processing and computer vision. Traditionally, these closely related fields were developed independently. However, the use of the level set method and variational methods in recent years started to bring all these fields together. One example is the TV-inpainting model [7]. We can perform inpainting in a desired domain while applying the ROF model [10] to remove noise from the rest of the domain using only one energy functional.

There are two interesting recent developments about the connection between different fields in image processing. We will discuss later in this paper how they link different fields in an interesting way. The first development concerns the impulse-noise removal method and the variational method for image regularization. In two recent papers by Chan, Nikolova et al. [3,4], a two-phase method was proposed to remove impulse-type noise. For a true image $u^*(x)$ and an observed image $f(x)$ defined in a domain Ω, impulse-type noise is defined by

$$f(x) = \begin{cases} r(x) & \text{with probability } r_0 \\ u^*(x) & \text{with probability } (1 - r_0). \end{cases} \tag{1}$$

As an example, for the so-called salt-and-pepper noise, $r(x)$ is simply the maximum or the minimum of the image intensity. The main idea in those papers is to separate the denoising process into a noise detection phase and a noise removal phase. In the first stage,

N. Paragios et al. (Eds.): VLSM 2005, LNCS 3752, pp. 149–160, 2005.

a median-type filter is applied to the observed image to detect the possible locations of the impulse noise. Then in the second phase, instead of replacing the intensity at all locations by the median intensity around a certain neighborhood, an L^1-regularization method is applied only to those locations reported in the first phase while keeping the other pixels unchanged. The resulting method was shown to be able to remove salt-and-pepper noise efficiently even at a very high noise level (for example $r_0 = 0.75$). The main reason for its success is that this method both retains those pixels that are unlikely to be polluted and maintains sharp edges in the whole image.

Another interesting development is a new model that uses a variational method for image segmentation [5,1]. The idea of the model is to minimize an energy functional consisting of a weighted TV-norm with an L^1-fidelity term. For segmentation of a binary image, the papers showed an equivalence between the new energy functional and that of the active contour model. This relation can be used to overcome the problem of the active contour model in which the energy function is not convex. Very often the snake will be trapped into a **local** minimum thus giving unsatisfactory segmentation results. The link between these two energies, as demonstrated in the above papers, provides a convenient way to determine the **global** minimizer of the active contour energy.

In this paper, we will combine these two recent advances in image processing. This provides an efficient way to bring image denoising (for both impulse-type and Gaussian-type noise), image segmentation, and image inpainting together. For low dimensional images, one could of course solve each of these tasks individually. It is computationally inexpensive to do so. As a result, it might not be obvious to see any advantage of having such a combined strategy for one or two dimensional images. However, for higher dimensional images, dealing with these problems separately might not be practical. The main difficulty concerns the computational time. With over 10^6 voxels, it could take days to analysis only one 3D scan. Therefore, by solving only one variational problem which performs all these tasks simultaneously, we could significantly speed up the computational time.

The rest of the paper is organized as follows. In Section 2, we will briefly review the denoising model by Chan, Nikolova et al. [3,4] and also restate the link between the active contour model and the ROF model as in [1]. A new model will be given in Section 3. Section 4 contains some details about the numerical implementation. Some numerical results are given in Section 5.

2 Two Recent Developments

2.1 A Two-Phase Method to Remove Impulse-Type Noise

Unlike the usual way to denoise impulse noise by applying the median-type filter to the image and replacing the image intensity everywhere, the idea in [3,4] is to separate the denoising processing into a noise-detection phase and a noise-removal phase. Mathematically, the first phase can be formulated as determining a noise candidate set $\mathcal{N} = \{x \in \Omega : f(x) \neq f_{\text{MF}}(x)\}$ in which Ω is the image domain, $f(x)$ is the observed image intensity at the pixel x, and $f_{\text{MF}}(x)$ is the intensity at x after applying a median-type filter, such as the classical median filter or the adaptive median filter. In the second phase, the following functional is minimized

$$F|_{\mathcal{N}}(u) = \int_{\mathcal{N}} \left\{ |u(x) - f(x)| + \frac{\beta}{2}[S_1(u) + S_2(u)] \right\} \tag{2}$$

where

$$S_1 = \int_{\mathcal{V}(x)\cap(\Omega\setminus\mathcal{N})} 2\,\phi[u(x) - f(y)]dy \text{ and } S_2 = \int_{\mathcal{V}(x)\cap\mathcal{N}} \phi[u(x) - u(y)]dy\,, \tag{3}$$

$\mathcal{V}(x)$ is the neighborhood centered at x and ϕ is an edge-preserving potential function. As seen in [3,4], one possible choice for ϕ is $\phi(t) = \sqrt{t^2 + \epsilon^2}$ with a small constant ϵ. The first term in the curve-bracket is an L^1-fidelity term. The terms in the square-bracket can be interpreted as an approximation of the total variation of u.

In the simple case when the noise can be separated accurately in the first step, the fidelity term is not important. This whole algorithm can then be simplified and the resulting variational method is the same as an image inpainting algorithm. For example, if ROF [10] or L^2 squared fidelity is used instead, we arrive at the TV-inpainting of [7]. That is, given an observed image f, one minimizes the following energy

$$E_1(u) = \int_\Omega |\nabla u| + \frac{1}{2}\int_\Omega \lambda(x)|u - f|^2 \tag{4}$$

where $\lambda(x) = 0$ if $f(x) = f_{\text{MF}}(x)$ and $\lambda(x) = \lambda_\infty \simeq \infty$ otherwise.

The idea of using a piecewise constant $\lambda(x)$ in TV-inpainting is not new [7]. However, it is interesting to see here the relationship between impulse-type noise removal and image inpainting by using a $\lambda(x)$ determined by a median-type filter.

2.2 Global Minimizer of the Active Contour Model

In the classical active contour model, the initial guess of the segmented image plays a very important role. We show in Figure 1 some minimizers of the active contour model. As we can see, different initial conditions in the evolution will give different segmented region. More importantly, none of these results corresponds to the *true* segmented results, i.e. curves which separate all regions with different intensities in the whole image. One reason for these unsatisfactory results is that the minimization problem of the active contour is not convex, and therefore it is very likely that the energy minimization could be trapped into a **local** minimium.

Recently, a few algorithms were proposed [5,1] to determine the **global** minimizers of some image segmentation models. In particular, an algorithm to determine the global minimizer of the active contour model based on the ROF model was given in [1]. The idea is to modify the ROF energy

$$E_{\text{ROF}}(u, \lambda) = \int_\Omega |\nabla u| + \frac{\lambda}{2}\int_\Omega |u - f|^2 \tag{5}$$

by first replacing the TV-norm by a weighted TV-norm and then, more importantly, changing the measure in the fidelity term from the square of the L^2-norm to the L^1-norm. This gives

$$E_2(u, \lambda) = \int_\Omega \tilde{g}(f)|\nabla u| + \lambda\int_\Omega |u - f| \tag{6}$$

in which $\tilde{g}(f) = 1/(1 + \beta|\nabla f|^2)$.

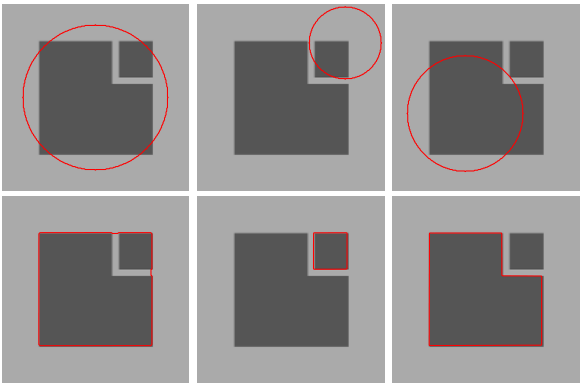

Fig. 1. Segmentation results using the active contour model. We show different initial configurations of the snake on the first row. The corresponding segmented results using these initial conditions are shown on the second row.

As pointed out in [1], if u is the characteristic function of a set Ω_C with boundary given by the curve C (i.e. $u = \mathbf{1}_{\Omega_C}$), the minimizer of the above energy E_2 is the same as the minimizer of the active contour energy

$$E_{\text{AC}}(C) = \int_C \tilde{g}(f)ds \tag{7}$$

with f approximated (in the sense of L^1) by a binary function of a region Ω_C.

Numerically, the minimization problem (6) is convex. This means that the method of gradient descent will converge to a unique minimizer, i.e. the global minimum of the energy function, independent of the initial condition. This equivalence is significant because by minimizing (6), one can determine the global minimium of the active contour model (7) avoiding both the danger of being trapped into any local minimum and the uncertainty in picking an initial configuration of the snake.

3 The New Energy

3.1 The Energy

Here, we propose a new model to combine those developments mentioned above. Given an observed image f, we minimize the energy

$$E(u) = \int_\Omega g(f)|\nabla u| + \int_\Omega \lambda(x)|u - f|\,. \tag{8}$$

This energy is similar to (6), except that $\lambda(x)$ is now changed into a function of the space variable and the weight in the weighted TV-norm is also modified. The function $\lambda(x)$ has the following properties.

$$\lambda(x) = \begin{cases} 0 & \text{TV-inpainting} \\ \lambda_0 & \text{Denoising} \\ \lambda_\infty \simeq \infty & u \text{ remains unchanged.} \end{cases} \tag{9}$$

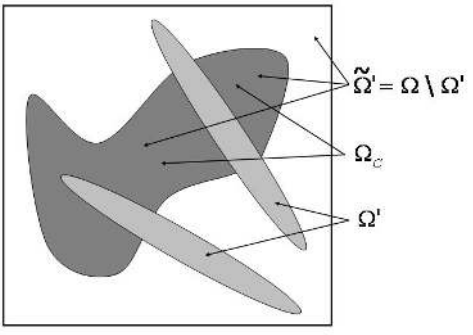

Fig. 2. Problem setting. Definition of the set Ω' (domain for inpainting), $\tilde{\Omega}'$ (compliment of Ω') and Ω_C (domain bounded by the curve C).

In the subdomain for image inpainting, $f(x)$ (and also $\tilde{g}(f)$) might not be known. We therefore simply set $g(f) = 1$, or $\beta = 0$. For the rest of the domain, we keep $g(f) = \tilde{g}(f)$.

Here we give some suggestions in picking such a function $\lambda(x)$ and also provide a variation in using the above minimization algorithm. For salt-and-pepper noise, the following $\lambda(x)$ works efficiently. We define $d(x)$ to be the difference in the intensities between the original image $f(x)$ and the modified image after applying the median-type filter $f_{\text{MF}}(x)$, i.e. $d(x) = |f(x) - f_{\text{MF}}(x)|$. Then one can set

$$\lambda_1(x) = \begin{cases} \lambda_\infty \text{ if } d(x) = 0 \text{ and } x \notin \Omega' \\ 0 \qquad\qquad \text{otherwise} \end{cases} \tag{10}$$

where $\Omega' \subset \Omega$ is a given subdomain for doing image inpainting and this subdomain can also be characterized by an user predefined mask function. This means that if x is in the inpainting domain Ω' or if the noise-detector detects that the image at x is polluted (therefore $f(x)$ will be different from the intensity after applying the median-type filter $f_{\text{MF}}(x)$), then the intensity at x will be modified by a TV-type regularization. Otherwise, the intensity at that location will remain unchanged.

If the impulse noise is random-valued instead, one can use a similar $\lambda(x)$

$$\lambda_2(x) = \begin{cases} \lambda_0 \text{ if } d(x) = 0 \text{ and } x \notin \Omega' \\ 0 \qquad\qquad \text{otherwise} \end{cases} \tag{11}$$

with $\lambda_0 << \lambda_\infty$.

For Gaussian-type noise, one can simply use

$$\lambda_3(x) = \begin{cases} \lambda_0 & \text{if } x \notin \Omega' \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

In the case when the type of noise is not known *a priori*, one can try to minimize (8) iteratively. More specifically, given the observed image $u_0 = f$, for $m = 1, \cdots, m_{\max}$, one iteratively minimizes

$$E(u_m) = \int_\Omega g(u_{m-1})|\nabla u_m| + \int_\Omega \lambda_4(x)|u_m - u_{m-1}| \tag{13}$$

with

$$\lambda_4(x) = \begin{cases} \lambda_0 \text{ if } d(x) \le d^* \text{ and } x \notin \Omega' \\ 0 \qquad\quad \text{otherwise} \end{cases} \tag{14}$$

where d^* is a threshold in the intensity difference function $d(x) \equiv |u_{m-1} - (u_{m-1})_{\text{MF}}|$.

3.2 The Link Between Active Contour for Segmentation, Denoising and TV-Inpainting

We explain here the relations between the minimization of the energy functional (8), the active contour model and the TV-inpainting model. Assuming $\Omega' = \{x \in \Omega : \lambda(x) = 0\}$ is the subdomain for inpainting (note that $\mathcal{N} \subset \Omega'$) and $\tilde{\Omega}' = \Omega \setminus \Omega'$, we have

$$E(u) = \int_{\tilde{\Omega}'} g(f)|\nabla u| + \int_{\tilde{\Omega}'} \lambda_0|u - f| + \int_{\Omega'} |\nabla u| = E^1(u) + E^2(u) \tag{15}$$

where

$$E^1(v) = \int_{\tilde{\Omega}'} g(f)|\nabla v| + \int_{\tilde{\Omega}'} \lambda_0|v - f| \text{ and } E^2(w) = \int_{\Omega'} |\nabla w| \tag{16}$$

with $v : \tilde{\Omega}' \to [u_{\min}, u_{\max}]$ and $w : \Omega' \to [u_{\min}, u_{\max}]$. So minimizing $E(u)$ is the same as $\min_v E^1(v) + \min_w E^2(w)$, and the minimizer of $E(u)$ is given by $u = \mathbf{1}_{\tilde{\Omega}'}(x) \cdot v + \mathbf{1}_{\Omega'}(x) \cdot w$ in which $\mathbf{1}_{\tilde{\Omega}'}$ is the characteristic function of the set $\tilde{\Omega}'$.

First we consider the energy $E^1(v)$. If Ω_C is a set in $\tilde{\Omega}'$ whose boundary is denoted by C and if the minimizer of $E^1(v)$ is given by $v = \mathbf{1}_{\Omega_C}$, then we have

$$E^1(v) = \int_{\tilde{\Omega}'} g(f)|\nabla \mathbf{1}_{\Omega_C}| + \int_{\tilde{\Omega}'} \lambda_0|\mathbf{1}_{\Omega_C} - f| = \int_C g(f)ds + \int_{\tilde{\Omega}'} \lambda_0|\mathbf{1}_{\Omega_C} - f|\,. \tag{17}$$

Therefore, minimizing $E^1(v)$ in the subdomain $\tilde{\Omega}'$ in the case of a binary observed image is equivalent to minimizing the active contour energy in $\tilde{\Omega}'$ while approximating f (in the L^1 sense) in $\tilde{\Omega}'$ by a binary function of Ω_C.

For the energy $E^2(w)$ defined in the complement, Ω', we have $E^2(w) = \int_{\Omega'} |\nabla w|$ and the boundary condition $w|_{\partial\Omega'} = v|_{\partial\Omega'}$ where v is the minimizer of $E^1(v)$. In the case when v is binary on $\partial\Omega'$, we have $w = \mathbf{1}_{\Omega_{C'}}$ again. This gives

$$E^2(w) = \int_{\Omega'} |\nabla \mathbf{1}_{\Omega_{C'}}| = \int_{C'} ds\,. \tag{18}$$

This implies that when the value of v on the boundary $\partial\Omega'$ is binary, minimizing $E^2(w)$ in Ω' is equivalent to $\min_{C'} \int_{C'} ds$ while the end points of C' are fixed on $\partial\Omega'$. Further analysis on the behavior of TV-inpainting can be found in [6].

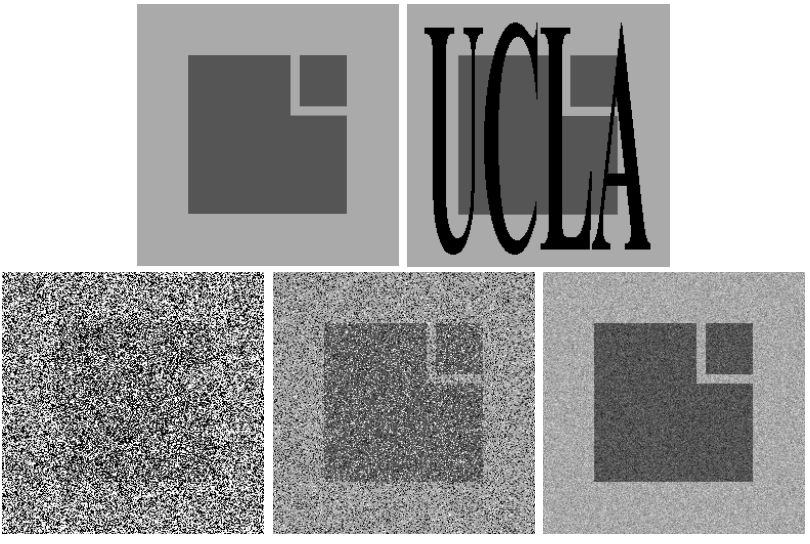

Fig. 3. (Example 1) (First row) The original true image and the corresponding image with an user defined mask (in black). (Second row) The original image with 75% salt-and-pepper noise, 50% random-valued impulse noise and additive Gaussian noise ($\sigma = 20$) respectively.

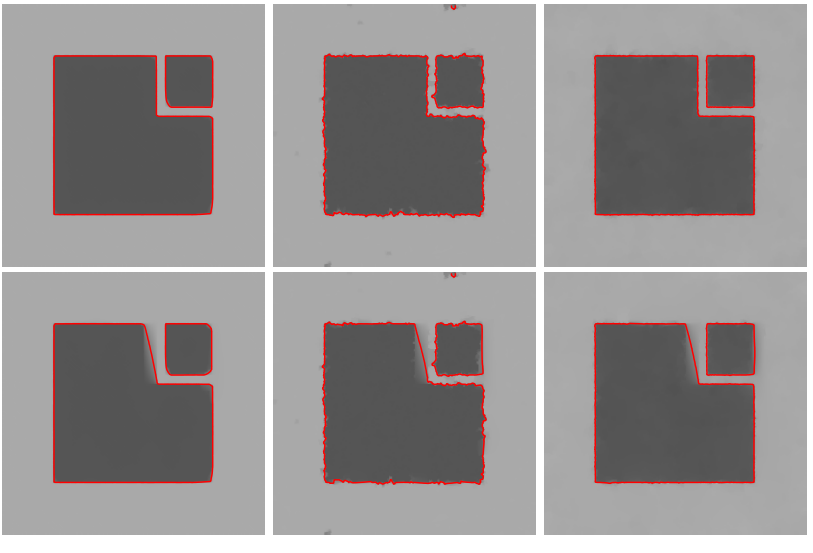

Fig. 4. (Example 1) The minimizer for the energy (8) **without** (first row) and **with** (second row) an extra mask. The left, middle and right columns show the denoising results with salt-and-pepper, random-valued impulse and addition Gaussian noise.

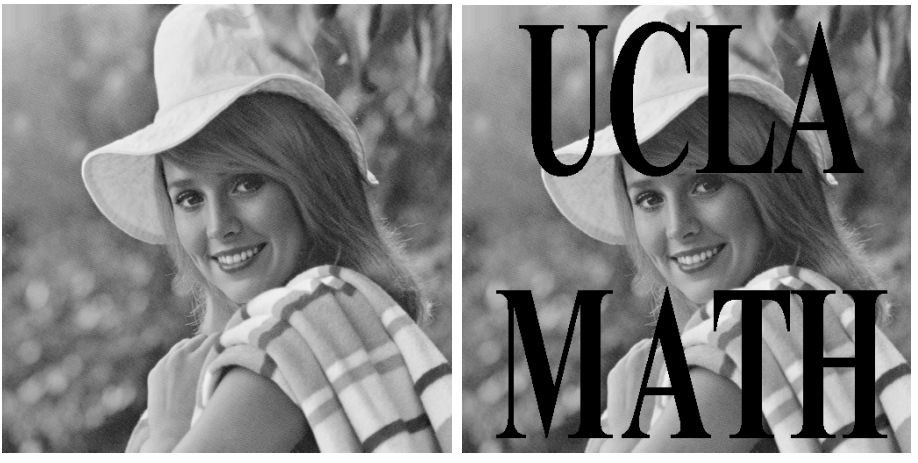

Fig. 5. (Example 2) The true image and the corresponding image with an user defined mask (in black)

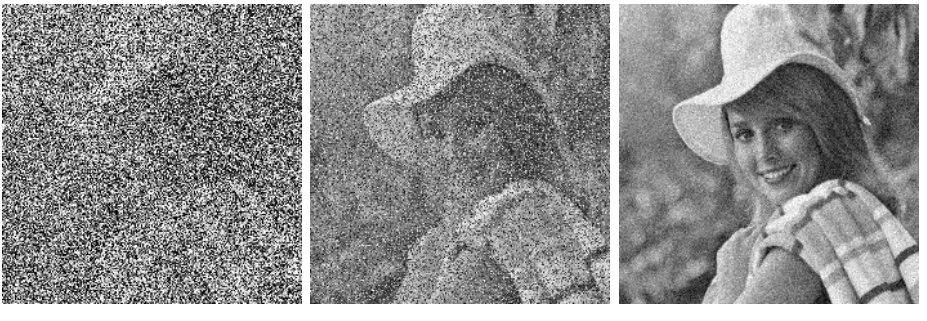

Fig. 6. (Example 2) The true image with 75% salt-and-pepper noise, 50% random-valued impulse noise and additive Gaussian noise ($\sigma = 20$) respectively

4 Numerical Method

To minimize the above energy, we use the method of gradient descent. The Euler-Lagrange equation of the energy functional (8) is given by

$$\frac{\partial u}{\partial t} = \nabla \cdot \left(\frac{g(x)}{|\nabla u|} \nabla u \right) - \lambda(x) \frac{u - f}{|u - f|} \,. \tag{19}$$

This equation can be solved numerically using the standard fully explicit scheme. In the current paper, however, we solve (19) using the Alternative Direction Explicit (ADE) technique.

5 Examples

In the following examples, we use $u^0(x, y) = 0$ as the initial condition for the Euler-Lagrange equation. Unlike the classical active contour/snake model, different initial

guesses used will give the **same global** minimizer of the segmentation model in the case of binary images.

5.1 Example 1

The true image used in this example has 256×256 pixels and is shown on the top left in Figure 3. The corresponding image with an user predefined mask is shown on the right on the same row. The black region is the domain Ω' where we want to perform TV-inpainting. The second row of Figure 3 shows the noisy versions of the true image.

The first row of Figure 4 shows the denoised image together with its segmentation with 75% salt-and-pepper noise, random-valued impulse noise and additive Gaussian noise. As we see from the figures, all noise is completely removed from the image. The red curves on the graphs are the boundaries of the segmented regions. Unlike the minimization of the active contour model, we can now easily reach the global minimum of the energy for these binary images regardless of the initial condition of the Euler-Lagrange equation. In the case of denoising together with image inpainting, the segmented results are shown on the second row of Figure 4. Again, all noise is completely removed from the image and we are able to fill in the missing part of the image using only one energy function.

5.2 Example 2

The true image of Elaine used in this example has 512×512 pixels and is shown on the left of Figure 5. On the right, we give the corresponding image with a predefined mask. 75% salt-and-pepper noise, 50% random-valued impulse noise and additive Gaussian noise with standard deviation $\sigma = 20$ are added to the original image and these observed images are shown in Figure 6. The minimizers of the energy functional (8) for these noisy images are shown in Figure 7. Since the original image is not binary, different level contours of the final figure will give different minimizers [1]. Using the case where salt-and-pepper noise and an external mask are added, we plot in Figure 8 the segmented results using the levels $u = 128$ and $u = 170$.

5.3 Example 3

Figures 9 shows the denoising results of a 3D brain MRI image. The number of voxels is $128 \times 256 \times 256$. The computational time is approximately 354 mins.

6 Conclusion

In this paper, we proposed combining the three most fundamental tasks in image processing - denoising, inpainting and segmentation. The idea is based on [5,1] but with the Lagrange multiplier, $\lambda(x)$, carefully chosen. At those locations where inpainting is required, we can set $\lambda(x)$ to be zero, which gives TV-inpainting. Indeed, TV-inpainting may not be the perfect model. For example, as mentioned in [7], TV-inpainting may not satisfy the so-called Connectivity Principle and may give unpleasant results to human vision. To improve this, one could modify the TV-regularization in the current

Fig. 7. (Example 2) The minimizer for the energy (8) **without** (first column) and **with** (second column) an extra mask. The top, middle and bottom rows show the denoising results with salt-and-pepper, random-valued impulse and addition Gaussian noise.

Fig. 8. (Example 2) Different segmented results with level (left) $u = 128$ and (right) $u = 170$ for the case with salt-and-pepper noise and an extra mask, i.e. from the top right picture in Figure 7

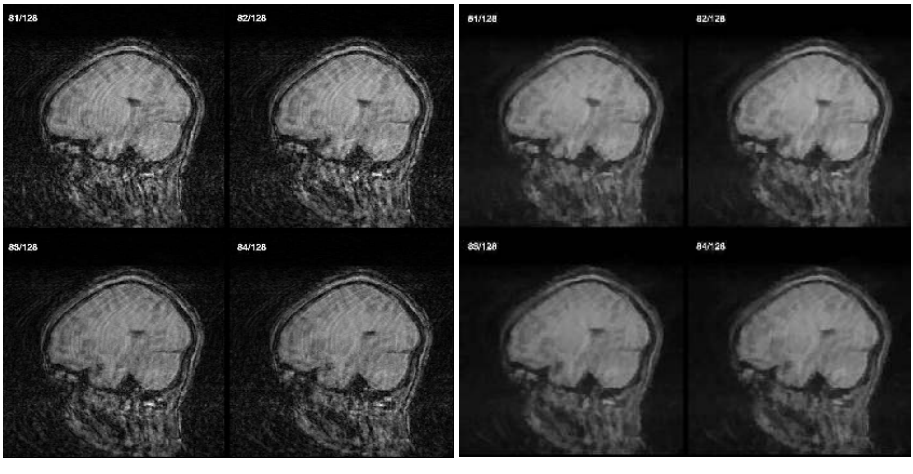

Fig. 9. (Example 3) (Left) A few slices of a noisy brain MRI image and (right) their corresponding denoised version

framework by a higher order interpolation model, for example, based on the Euler's Elastica [7]. As another drawback of the model, we indeed obtain multiple minimizers of the active contour model for non-binary images, as seen in Figure 8. This limitation is inherited from the original theory as described in [1].

Acknowledgment

This work was supported in part by NSF contract DMS-0312222 and NIH contract U54 RR021813.

References

1. X. Bresson, S. Esedoḡlu, P. Vandergheynst, J.P. Thiran, and S. Osher. Global minimizers of the active contour/snake model. *UCLA CAM Report (05-04)*, 2005.
2. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79, 1997.
3. R. Chan, C.-W. Ho, and M. Nikolova. Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization. *In preparation*, 2004.
4. R. Chan, C. Hu, and M. Nikolova. An iterative procedure for removing random-valued impuse noise. *In preparation*, 2004.
5. T.F. Chan, S. Esedoḡlu, and M. Nikolova. Algorithms for finding global minimizers of image segmentation and denoising models. *UCLA CAM Report (04-54)*, 2004.
6. T.F. Chan and S.H. Kang. Error analysis for image inpainting. *UCLA CAM Report (04-72)*, 2004.
7. T.F. Chan and J. Shen. Variational image inpainting. *Comm. Pure Appl. Math.*, 58(5):579–619, 2005.
8. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1996.
9. S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A.J. Yezzi. Conformal curvature flows: From phase transitions to active vision. *Archive for Rational Mechanics and Analysis*, 134:275–301, 1996.
10. L. Rudin, S.J. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.