UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

Reduced Non-Convex Functional Approximations for Image Restoration & Segmentation

Luminita Vese
Tony F. Chan

December 1997

CAM Report 97-56

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90095-1555

REDUCED NON-CONVEX FUNCTIONAL APPROXIMATIONS FOR IMAGE RESTORATION & SEGMENTATION*

LUMINITA VESE† AND TONY F. CHAN‡

Abstract. This paper is devoted to some approximations of functionals arising in image recovery and segmentation by some non-convex functionals. We show that the approximation of Ambrosio-Tortorelli for the Mumford-Shah functional and the approximation of Shah for a total variation minimization variant, consisting in solving a system formed by two equations with two unknowns, can be practically reduced to some non-convex approximations with only one equation and only one unknown. We also show how we can derive more general sequences of approximations. Our models can simultaneously be used for denoising and edge detection. Finally, using these various approximations, we present experimental results and comparisons for image restoration and segmentation, which show that our new models produce better images with sharper edges, due to the non-convexity. Also, the reduced algorithms are faster.

Key words. Variational problems, PDE's, non-convex functionals, Image processing, denoising, segmentation, numerical analysis.

AMS subject classifications. 49, 35, 65, 68.

1. Introduction. In this work, we deal with minimization of functionals depending on discontinuous functions. These functionals model various problems in image analysis, fracture mechanics, the theories of phase transitions and of liquid crystals, and other fields of mathematical physics. Here we are more interested in the application of such functionals in image recovery and segmentation. The general problem is: given an initial image u_0, to detect the relevant contours or, if this image has been corrupted by noise and/or blur, to recover it by preserving and extracting these contours (or edges). For an image-function u, the edges are discontinuities along curves. Therefore, the suitable function space to model images is the space of functions of bounded variation BV: only this space allows for these type of discontinuities, in comparison with the Sobolev spaces. But the difficulty is in general in the numerical approximation of the jumps set of the solution u, denoted by S_u (i.e. the edges).

An important model for image restoration and segmentation by variational methods is the Mumford-Shah functional [11], depending on two variables, the image-function u and a set S. The problem is: minimize

$$F^{MS}(u,S) = \int_{\Omega - S} (\alpha|\nabla u|^2 + \beta(u-u_0)^2)dxdy + \mathcal{H}^1(S),$$

where $\Omega \subset \mathbb{R}^2$ is open and bounded, $\mathcal{H}^1$ is the Hausdorff 1-dimensional measure in $\mathbb{R}^2$, S varies in the class of closed subsets of Ω, u varies in $C^1(\Omega - S)$ and $\alpha, \beta > 0$ are fixed positive parameters. The existence of minimizers is solved in a weak formulation of this problem ([1]), by setting $S = S_u$, the jumps set of u and minimizing only over

*This work is supported in part by ONR Contract N00014-96-1-0277 and NSF Contract DMS-9626755.

†UCLA, Mathematics Department, 405, Hilgard Ave, Los Angeles, Ca 90095-1555, U.S.A. (e-mail: lvese@math.ucla.edu)

‡UCLA, Mathematics Department, 405, Hilgard Ave, Los Angeles, Ca 90095-1555 (e-mail: chan@math.ucla.edu).

u. The associated functional is:

$$G^{MS}(u) = \int_\Omega \left(\alpha|\nabla u|^2 + \beta|u - u_0|^2\right)dxdy + \mathcal{H}^1(S_u),$$

where $u_0 \in L^\infty(\Omega)$. This functional has minimizers and moreover each minimizer belongs to $PC^1(\Omega)$, i.e. the space of piecewise C^1 functions ([1]).

From a numerical point of view, it is not easy to compute a minimizer for this functional, due to the term $\mathcal{H}^1(S_u)$, and to the fact that this functional is not a-priori the lower-semicontinuous relaxation of a more regular functional, defined for instance in $W^{1,1}(\Omega)$ or $H^1(\Omega)$ (it is not, for instance, the lower-semicontinuous relaxation of its restriction to these spaces).

We mention the merging regions method proposed to solve the Mumford-Shah functional, in the piecewise constant case, but which is not based on a PDE approach (see J.-M. Morel and S. Solimini [10], where several algorithms for segmentation by variational methods are presented).

Ambrosio and Tortorelli ([2], [3]) proposed two approximations for the Mumford-Shah functional ([11]), where they introduced an extra-variable v which approaches in some sense the jumps set of the solution u. In [13], J. Shah applied the same idea to approach a variant of the total variation minimization introduced by Rudin-Osher-Fatemi ([12]), also using a dual variable v as an edge-strength function. There approximations yields a coupled system of two equations and with two unknowns (u, v).

We have shown in a previous paper [8], that the approximations of Ambrosio-Tortorelli and Shah by functionals depending on two variables can be practically reduced to some non-convex approximations by functionals depending only on one variable u, the image function. v is expressed as a function of the gradient of u and the edges are still easily extracted from $v = v(|\nabla u|)$. Moreover, edges of images recovered by our non-convex functionals seems to be better preserved and the problems are simplified, because we do no more work with two variables (u, v), but only with u. In this way, we obtain faster algorithms.

In this paper we want to analyze further these reduced approximations and especially their relations with the Ambrosio-Tortorelli and Shah approximations. Exploiting these connections, we will propose other approximations and we show how we can derive more general approximations in one variable (still being able to extract the edges), or approximations in two variables generalizing the models of Ambrosio-Tortorelli and Shah. We will also study the numerical asymptotical behavior of our new approximations compared with the total variation minimization and show similarities with the theoretical result of Ambrosio-Tortorelli. Finally, we present several numerical results on images and compare further these different approximations.

We note that our work is also related to works by G. Aubert, M. Barlaud, L. Blanc-Féraud and S. Teboul ([5] and [14]). Their method is still based on an approximation by a coupled system with two equations and two unknowns. We want also to mention B. Bourdin [7], were the finite element method is used for the numerical implementation of the Ambrosio-Tortorelli approximation [3].

The outline of the paper is as follows. In the next section, we recall the approximations proposed by Ambrosio-Tortorelli and Shah, together with our reduced non-convex approximations, introduced in [8]. In Section 3, we will show more connections between these reduced approximations and the initial approximations of Ambrosio-Tortorelli, Shah, and we will derive more general approximations. Then, we propose

another approximation, following the same idea. By all these new approximations, we are able to reconstruct the image function u, in parallel with the edge detection.

Finally, the last section is devoted to several numerical results and comparisons applied to denoising and edge detection. We show that the obtained results by the new models are better and the locations of edges are well-preserved. Moreover, the algorithms are faster. In this last section we will also show the connections and comparisons with the total variation minimization, exploiting the numerical asymptotical behavior of the approximation.

2. Reduced non-convex functional approximations. In the approximations of Ambrosio-Tortorelli and Shah, the idea was to introduce a dual variable v which will "approach" passing to the limit, in some sense, the jumps set S_u (i.e. the edges) of the solution u.

Ambrosio and Tortorelli ([2],[3]) proposed two approximations for the Mumford-Shah functional. We will refer in this paper only to the second approximation ([3]), which is simpler than the first ([2]).

This approximation of Ambrosio-Tortorelli for G^{MS} is ([3]):

$$G_\rho^{AT}(u,v)=\int_\Omega\Big[\rho|\nabla v|^2+\alpha\Big(v^2|\nabla u|^2+\frac{(v-1)^2}{4\alpha\rho}\Big)+\beta|u-u_0|^2\Big]dxdy, \tag{2.1}$$

and they showed that if $w_\rho=(u_\rho,v_\rho)$ minimizes G_ρ^{AT}, then (passing to subsequences) u_ρ is an approximation of u, a minimizer of G^{MS}, and v_ρ goes to 1, as $\rho\to 0$, in the $L^2(\Omega)$-topology (i.e. $\int_\Omega|u_\rho-u|^2dxdy\to 0$ and $\int_\Omega|v_\rho-1|^2dxdy\to 0$, as $\rho\to 0$). Here, v_ρ is different from one (and less then one) only in a small neighborhood of S_u, which shrinks as $\rho\to 0^+$.

The system associated to the minimization of $G_\rho^{AT}(u,v)$ is:

$$\begin{cases}\beta u-\alpha\nabla(v^2\nabla u)=\beta u_0\\ -\triangle v+\frac{1+4\alpha\rho|\nabla u|^2}{4\rho^2}\Big(v-\frac{1}{1+4\alpha\rho|\nabla u|^2}\Big)=0.\end{cases} \tag{2.2}$$

We can see from the second equation of (2.2) that v is in fact a smoothing of $\frac{1}{1+4\alpha\rho|\nabla u|^2}$.

We want to reduce this system to only one equation with only one unknown, while still being able to extract the edges. We see that we can not simplify further or drop the first equation of (2.2) in u: the parameter α has to be strictly positive, because first of all, we have to regularize u to remove noise.

Our main new idea is to drop the term $-\triangle v$ from the second equation in v of (2.2). Then this equation in v has the explicit solution $v=v_\rho(|\nabla u|)=\frac{1}{1+4\alpha\rho|\nabla u|^2}$. Lets replace it in $G_\rho^{AT}(u,v)$ (i.e. without considering the regularizing term $\int_\Omega\rho|\nabla v|^2$ in $G_\rho^{AT}(u,v)$). Then we obtain, after computation, the functional:

$$\begin{aligned}G_\rho^{at}(u)&=\int_\Omega\Big(\alpha\frac{|\nabla u|^2}{1+4\alpha\rho|\nabla u|^2}+\beta|u-u_0|^2\Big)dxdy\\&=\int_\Omega\Big\{\alpha\Big[\Big(\frac{1}{1+4\alpha\rho|\nabla u|^2}\Big)^2|\nabla u|^2+\frac{1}{4\alpha\rho}\Big(\frac{1}{1+4\alpha\rho|\nabla u|^2}-1\Big)^2\Big]+\beta|u-u_0|^2\Big\}.\end{aligned}$$

In this way, the following new problem:

$$\begin{cases}\inf_u\Big\{\ G_\rho^{at}(u)\ \Big\} & \text{(restoration for } u)\\ v_\rho(|\nabla u|)=\frac{1}{1+4\alpha\rho|\nabla u|^2} & \text{(edge-detector)}\end{cases} \tag{2.3}$$

is our reduced model for the Ambrosio-Tortorelli approximation (2.1). Then we no more have to solve two coupled PDE's, but only one, from the minimization of $G_\rho^{at}(u)$, v being expressed as an explicit function of the gradient $|\nabla u|$ of u.

Let us now consider the following functional, introduced by J. Shah ([13]):

$$G^S(u) = \int_\Omega \Big(\alpha|\nabla u| + \beta|u - u_0|\Big)dxdy + \int_{S_u} \frac{\alpha|u^+ - u^-|}{1 + \alpha|u^+ - u^-|} d\mathcal{H}^1,$$

where $(u^+ - u^-)$ represents the jump of u across S_u.

J. Shah proposed the following approximation (following the idea of Ambrosio-Tortorelli), as $\rho \to 0$, for $G^S(u)$:

$$G_\rho^S(u,v) = \int_\Omega \Big[\rho|\nabla v|^2 + \alpha\Big(v^2|\nabla u| + \frac{(v-1)^2}{4\alpha\rho}\Big) + \beta|u - u_0|\Big]dxdy.$$

The system associated to the minimization of $G_\rho^S(u,v)$ is:

$$\text{(2.4)} \qquad \begin{cases} \beta\frac{u-u_0}{|u-u_0|} - \alpha\Big(v^2\frac{\nabla u}{|\nabla u|}\Big) = 0, \\ -\triangle v + \frac{1+4\alpha\rho|\nabla u|}{4\rho^2}\Big(v - \frac{1}{1+4\alpha\rho|\nabla u|}\Big) = 0, \end{cases}$$

and again v can be seen to be a smoothing of $\frac{1}{1+4\alpha\rho|\nabla u|}$. Let us drop the term $-\triangle v$ in the second equation of (2.4). Then we obtain $v = v_\rho(|\nabla u|) = \frac{1}{1+4\alpha\rho|\nabla u|}$ and we replace it in $G_\rho^S(u,v)$ (i.e. without the regularizing term $\int_\Omega \rho|\nabla v|^2$). We obtain the functional:

$$\begin{aligned} G_\rho^s(u) &= \int_\Omega \Big(\alpha\frac{|\nabla u|}{1 + 4\alpha\rho|\nabla u|} + \beta|u - u_0|\Big)dxdy \\ &= \int_\Omega \Big\{\alpha\Big[\Big(\frac{1}{1 + 4\alpha\rho|\nabla u|}\Big)^2|\nabla u| + \frac{1}{4\alpha\rho}\Big(\frac{1}{1 + 4\alpha\rho|\nabla u|} - 1\Big)^2\Big] + \beta|u - u_0|\Big\}. \end{aligned}$$

Again, as before, the following new problem:

$$\text{(2.5)} \qquad \begin{cases} \inf_u \Big\{ G_\rho^s(u) \Big\} & \text{(restoration for } u\text{)} \\ v_\rho(|\nabla u|) = \frac{1}{1+4\alpha\rho|\nabla u|} & \text{(edge-detector)} \end{cases}$$

is our reduced model for the Shah approximation, to only one PDE with one unknown u, v being expressed as a function of $|\nabla u|$.

Analogous to these new approximations, we may consider a more general model, as follows: first minimize

$$\text{(2.6)} \qquad G_\rho(u) = \int_\Omega \Big(\alpha\frac{|\nabla u|^p}{1 + 4\alpha\rho|\nabla u|^p} + \beta|u - u_0|^p\Big)dxdy,$$

with for instance $p \geq 1$, to restore u, and then compute v by:

$$v(|\nabla u|) = \frac{1}{1 + 4\alpha\rho|\nabla u|^p}$$

as an edge-detector. We obtain from (2.6) G_ρ^{at} for $p = 2$ and G_ρ^s for $p = 1$.

Remark.

(*i*) In our experimental results, we use the following functional, still denoted G_ρ and obtained by replacing $(u-u_0)^p$ in (2.6) by $(u-u_0)^2$, for all $p>0$:

$$G_\rho(u)=\int_\Omega\Big(\alpha\frac{|\nabla u|^p}{1+4\alpha\rho|\nabla u|^p}+\beta|u-u_0|^2\Big)dxdy. \tag{2.7}$$

On the other hand, J. Shah, to solve the system (2.4), used the associated evolutionary system, which is:

$$\begin{cases}\frac{\partial u}{\partial t}+\beta\frac{u-u_0}{|u-u_0|}-\alpha\Big(v^2\frac{\nabla u}{|\nabla u|}\Big)=0,\\ \frac{\partial v}{\partial t}-\triangle v+\frac{1+4\alpha\rho|\nabla u|}{4\rho^2}\Big(v-\frac{1}{1+4\alpha\rho|\nabla u|}\Big)=0,\end{cases} \tag{2.8}$$

with $u(0)=u_0$ and for $v(0)$ an approximation. The problem (2.8) will be called in this paper "the initial Shah model". We will compare this model also with the following (called here "the modified Shah model"), obtained by replacing the L^1-norm of $(u-u_0)$ in G_ρ^S by its L^2-norm. The associated stationary system is:

$$\begin{cases}2\beta u-\alpha\Big(v^2\frac{\nabla u}{|\nabla u|}\Big)=2\beta u_0,\\ -\triangle v+\frac{1+4\alpha\rho|\nabla u|}{4\rho^2}\Big(v-\frac{1}{1+4\alpha\rho|\nabla u|}\Big)=0.\end{cases} \tag{2.9}$$

(*ii*) In order to see how the problem changes when we do not consider the regularizing term $\int_\Omega\rho|\nabla v|^2$ in the functionals, which is equivalent to drop the Laplacian of v in the associated systems, let us write the second equation from (2.2) in the form:

$$v^{AT}=\frac{1}{1+4\alpha\rho|\nabla u|^2}+\frac{4\rho^2}{1+4\alpha\rho|\nabla u|^2}\triangle v^{AT}.$$

We see that the difference between our edge-detector $v=\frac{1}{1+4\alpha\rho|\nabla u|^2}$ and v^{AT} is only by a small amount of diffusion. This diffusion is function of the coefficient $\frac{4\rho^2}{1+4\alpha\rho|\nabla u|^2}$, which is very small because $\rho\to 0$. In other words, in our reduced models, we do not have an explicit diffusion on v, this being regularized implicitly, because $v=v(|\nabla u|)$ and u is regularized. However, the problems are practically simplified, because we have only one equation in u and v is computed only one time, in the end. Moreover, the numerical results are better.

(*iii*) Let us denote by ϕ the function $z\mapsto\frac{z^p}{1+4\alpha\rho z^p}$. Then G_ρ can be written as:

$$G_\rho(u)=\int_\Omega\Big(\alpha\phi(|\nabla u|)+\beta|u-u_0|^2\Big)dxdy, \tag{2.10}$$

and we will see in the numerical results that $v_\rho(|\nabla u|)=\frac{1}{1+4\alpha\rho|\nabla u|^p}$ may be used as an edge-strength function or edge detector, and moreover, just like in the theoretical result of Ambrosio-Tortorelli, we will have "numerically" that $v_\rho\to 1$ as $\rho\to 0$, except in a small neighborhood of S_u, but which shrinks as $\rho\to 0$. Note that the function ϕ is non-convex. For instance for $p=1$, G_ρ^s as $\rho\to 0$ is formally a non-convex approximation of the total variation minimization ([12]). By minimizing non-convex functionals (as for instance in [15]) in image recovery, the edges are sharper and better preserved than with convex functionals, but unfortunately the existence and uniqueness is an open problem. In this paper, we propose in fact a compromise between the minimization of a convex functional (as for instance the TV minimization, so the existence

and uniqueness are solved) and the minimization of a non-convex functional (when the edges are better preserved, but without knowing something about the existence and uniqueness of a minimizer): we approach the convex functional by a sequence of non-convex functionals (see Fig. 2.1 for $p = 1$ and $p = 2$).

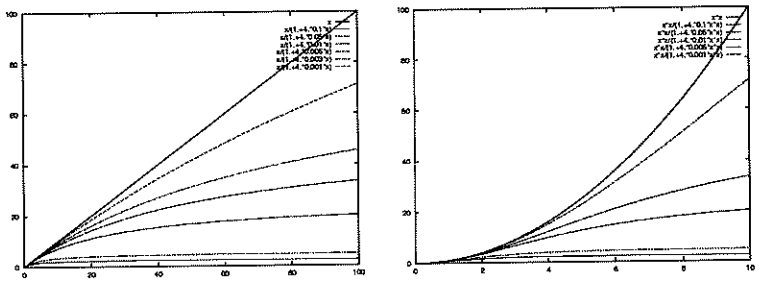

FIG. 2.1. *We illustrate how* $|\nabla u|$ *to the left (and* $|\nabla u|^2$ *to the right) is approximated by the non-convex function* $\frac{|\nabla u|}{1+\rho|\nabla u|}$ *(by* $(\frac{|\nabla u|^2}{1+\rho|\nabla u|^2}$ *respectively) as* $\rho \to 0$.

(*iv*) We will explain here why the edges can be better preserved with a non-convex (and sub-linear) functional like $\int_\Omega \phi(|\nabla u|)$, rather than with a convex functional (like TV minimization or Mumford-Shah functional). Let us write the Euler equation associated to the minimization of $G_\rho(u)$:

$$2\beta u - \alpha \nabla \cdot \Big(\frac{\phi^{'}(|\nabla u|)}{|\nabla u|} \nabla u \Big) = 2\beta u_0. \tag{2.11}$$

Let us define for each point (x, y) where $|\nabla u(x, y)| \neq 0$, the vectors $\vec{N}(x, y) = \frac{\nabla u(x,y)}{|\nabla u(x,y)|}$ in the gradient direction and $\vec{T}(x, y)$ in the orthogonal direction to $\vec{N}(x, y)$. We represent in Fig. 2.2 these vectors with the level curve of u passing through (x, y).

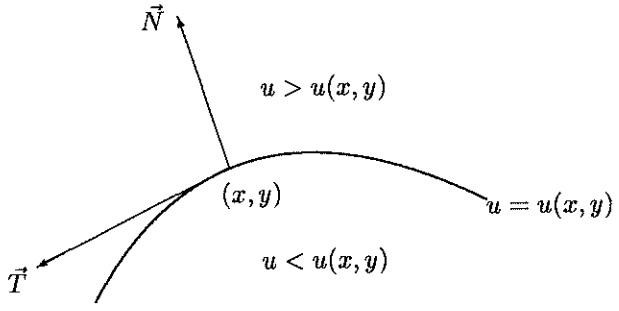

FIG. 2.2. *A contour C separating two homogeneous regions.*

With the usual notations $u_x, u_y, u_{xx}, \ldots$ for the first and second partial derivatives of u, and by formally developing the divergence operator in (2.11), we obtain:

$$2\beta u - \alpha \Big(\frac{\phi^{'}(|\nabla u|)}{|\nabla u|} \Big) u_{TT} - \alpha \phi^{''}(|\nabla u|) u_{NN} = 2\beta u_0, \tag{2.12}$$

where we have denoted by u_{TT} and u_{NN} the second derivatives of u in the directions

$\vec{T}(x,y)$ and $\vec{N}(x,y)$ respectively:

$$u_{TT} = \frac{1}{|\nabla u|^2}(u_x^2 u_{yy} - 2u_x u_y u_{xy} + u_y^2 u_{xx}),\ u_{NN} = \frac{1}{|\nabla u|^2}(u_x^2 u_{xx} + 2u_x u_y u_{xy} + u_y^2 u_{yy}).$$

In a neighborhood of a contour $\mathcal{C}$ (see Fig. 2), where the image presents strong gradients, in (2.12), the term $-\alpha\phi'' u_{NN}$ will have an anti-diffusion effect due to the non-convexity of ϕ, and this will reinforce the edges (this term is an anti-diffusion in the orthogonal direction to the edge, i.e. a diffusion operator with the "wrong" sign). We note that all functions $(z \mapsto \phi(z),\ z \geq 0)$ for $p \geq 1$, are non-convex for "large" z.

3. General functional approximations. In this section, we want to generalize our reduced approximations presented in the previous section, by considering a more general function ϕ, still being able to extract the edges. This can be done using the relations between our reduced approximations and the approximations of Ambrosio-Tortorelli and Shah.

Let us first make some notations and denote by ϕ_ρ and ψ_ρ the following functions:

$$\phi_\rho(t) = \frac{t}{1+4\alpha\rho t},\quad \psi_\rho(v) = \frac{(v-1)^2}{4\alpha\rho}.$$

Then, the general model (2.7) can be written as:

$$G_\rho(u) = \int_\Omega \left(\alpha\phi_\rho(|\nabla u|^p) + \beta|u-u_0|^2\right) = \int_\Omega \left(\alpha(v_\rho^2|\nabla u|^p + \psi_\rho(v_\rho)) + \beta|u-u_0|^2\right),$$

where

$$v_\rho = v_\rho(|\nabla u|) = \frac{1}{1+4\alpha\rho|\nabla u|^p}.$$

3.1. The relation between ϕ_ρ, v_ρ and ψ_ρ. It can easily be verified (and this is natural from the construction of the reduced models) that, for all $p \geq 0$, we have:

$$\phi_\rho(|t|^p) = \inf_v (v^2|t|^p + \psi_\rho(v)). \tag{3.1}$$

Moreover, the infimum is reached for $v(|t|) = \sqrt{\phi'_\rho(|t|^p)} = \frac{1}{1+4\alpha\rho|t|^p}$ (compare with $v_\rho(|\nabla u|) = \frac{1}{1+4\alpha\rho|\nabla u|^p}$ derived before).

We can say that our approximations depending on only one variable are new formulations of the Ambrosio-Tortorelli and Shah approximations with two variables. Of-course, we do not say that our new model is equivalent with the approximations of Ambrosio-Tortorelli ($p=2$) and Shah ($p=1$), or even with the Mumford-Shah functional (again $p=2$). We already pointed out that $\phi_\rho(|t|^p)$ is a non-convex approximation of $|t|^p$ as $\rho \to 0$. We note that this approximation by a non-convex functional of this type can be related to a problem arising in phase transition (see [6]).

3.2. General approximations by non-convex functionals. Using the idea of the previous subsection and considering general functions ϕ instead of $\phi(t) = \frac{t}{1+4\alpha\rho t}$ but with the same properties, we will construct more general sequences of approximations. Let us first present the following result:

THEOREM 3.1. *Let $\phi : [0,+\infty[\to [0,+\infty[$ be a positive C^2 function satisfying:*
(i) $\phi(0) = 0,\ \ 0 < \lim_{t\to 0}\phi'(t) = \phi'(0) < +\infty$;

(ii) $\lim_{t\to\infty}\phi'(t)=0$;
(iii) ϕ' *is continuous and strictly decreasing on* $[0,+\infty[$;
(iv) $t\phi'(t)+\phi(t)$ *is increasing on* $[0,+\infty[$.
Then there is a convex and decreasing function $\psi:]0,\sqrt{\phi'(0)}]\to\mathbb{R}$ *such that:*

$$\phi(t)=\inf_v(v^2t+\psi(v)),\quad\forall\, t\geq 0. \tag{3.2}$$

Moreover, for a given $t\geq 0$, *the infimum will be reached for*

$$v_t=\sqrt{\phi'(t)},$$

and the function ψ *is given by one of the following expressions:*

$$\psi(v)=\phi\Big((\phi')^{-1}(v^2)\Big)-v^2\cdot(\phi')^{-1}(v^2), \tag{3.3}$$

$$\psi(v)=-\phi^*(v^2)=-\sup_t(v^2t-\phi(t))=\inf_t(-v^2t+\phi(t)), \tag{3.4}$$

where ϕ^* *is the Legendre-Fenchel transform of* ϕ. *Defining the functions* Φ *on* $L^1(\Omega)$ *and* Ψ *on* $L^\infty(\Omega)$ *by*

$$\Phi(u)=\int_\Omega\phi(|u(x,y)|)dxdy,$$
$$\Psi(v)=\int_\Omega\psi(v)dxdy,$$

we have the same relations between Φ *and* Ψ:

$$\Phi(u)=\inf_v\Big(\int_\Omega v^2u dxdy+\Psi(v)\Big),$$
$$\Psi(v)=\inf_u\Big(-\int_\Omega v^2u dxdy+\Phi(u)\Big).$$

Proof. Let $t_0\geq 0$ be arbitrary fixed and let $v_0=\sqrt{\phi'(t_0)}$. Because the function ϕ' is invertible from the assumptions, then $t_0=(\phi')^{-1}(v_0^2)$. The function ψ is well-defined by the formula in (3.3) and we have:

$$\psi(v_0)=\phi(t_0)-v^2t_0,$$

or

$$\phi(t_0)=v_0^2t_0+\psi(v_0).$$

Now we will show that:

$$\phi(t_0)=v_0^2t_0+\psi(v_0)\leq v^2t_0+\psi(v),\ \forall v\in]0,\sqrt{\phi'(0)}],$$

or

$$\phi(t_0)\leq v^2t+\psi(v)+v^2t_0-v^2t,\ \forall v\in]0,\sqrt{\phi'(0)}],$$

which is equivalent to showing:

$$\phi(t_0)\leq\phi(t)+\phi'(t)(t_0-t),\ \forall t\geq 0$$

(because, for every $v \in]0, \sqrt{\phi'(0)}]$, there is a unique $t \geq 0$ s.t. $v^2 = \phi'(t)$ and then $v^2 t + \psi(v) = \phi(t)$).

We have

$$\phi(t) - \phi(t_0) + \phi'(t)(t_0 - t) = (t - t_0)(\phi'(c) - \phi'(t)), \ \forall t \geq 0,$$

for some c between t_0 and t, depending on t, and we can easily verify the inequality:

$$0 \leq (t - t_0)(\phi'(c) - \phi'(t)), \ \forall t \geq 0.$$

Finally we deduce that:

$$\phi(t_0) = \inf_v (v^2 t_0 + \psi(v)),$$

with ψ given by (3.3) and the infimum being reached for $v_0 = \sqrt{\phi'(t_0)}$.

It can easily be checked that $\psi'(v) = -2v(\phi')^{-1}(v^2)$, so ψ is decreasing. In order to show that ψ is convex, we compute ψ'' and we use (iv):

$$\psi''(v) = -2(\phi')^{-1}(v^2) - \frac{4v^2}{\phi''((\phi')^{-1}(v^2))} = -2t - \frac{4\phi'(t)}{\phi''(t)} \geq 0,$$

where $t = (\phi')^{-1}(v^2)$.

The assertions related to the functions Φ and Ψ can easily be verified. We note that $Dom(\phi^*) =]0, \phi'(0)]$. □

Example. Let consider for instance $\varphi_\rho(t) = \frac{1}{\rho}\log(1 + \rho t)$. Then $\varphi_\rho(|t|) \to |t|$ as $\rho \to 0$. This function satisfies the assumptions of Theorem 3.1. Then there is a convex and decreasing function ξ_ρ s.t.

$$\varphi_\rho(|t|^p) = \frac{1}{\rho}\log(1 + \rho|t|^p) = \inf_v (v^2|t|^p + \xi_\rho(v)),$$

the infimum being reached for $v(|t|) = \sqrt{\varphi'(|t|^p)} = \sqrt{\frac{1}{1+\rho|t|^p}}$, and ξ_ρ is given by: $\xi_\rho(v) = \frac{1}{\rho}(v^2 - 2\log v - 1)$.

Then we obtain a new model (or a new approximation): minimize

$$E_\rho(u) = \int_\Omega \left(\alpha\varphi_\rho(|\nabla u|^p) + \beta|u - u_0|^2\right) dxdy, \tag{3.5}$$

and use

$$v_\rho(|\nabla u|) = \sqrt{\frac{1}{1 + \rho|\nabla u|^p}}$$

as an edge-detector, whith for instance $p \geq 1$.

We note that even if the functions $\phi_\rho(t) = \frac{t}{1+4\alpha\rho t}$ and $\varphi_\rho(t) = \frac{1}{\rho}\log(1+\rho t)$ satisfy both the properties of Theorem 3.1, there are also some differences: for instance, ϕ_ρ has an horizontal asymptote to infinity, while φ_ρ growths (sub-linearly) to infinity. We will compare numerically these two functions, for different values of p, with $1 \leq p \leq 2$ (we expect to obtain better results for u using ϕ_ρ to reconstruct piecewise-constant images, but for images which are no longer piecewise constant, the second function φ has to produce better results for u).

In the last section, devoted to several experimental results, we will see that we can use these reduced-generalized models to reconstruct the image function u, in parallel with edge-detection, computing in the end the function v. Like in the Ambrosio-Tortorelli approximation, v is formally different from $\sqrt{\phi'(0)}$ (and less than $\sqrt{\phi'(0)}$) only in a small neighborhood of the jumps set S_u (the edges) and this neighborhood shrinks as $\rho \to 0$. We will compare the behavior of the results for ϕ, $p = 1$ and $\rho \to 0$ with the total variation minimization ($\rho = 0$). We will see that we obtain "numerically" the complete convergence of u_ρ and v_ρ to respectively u_{TV} and 1, as $\rho \to 0$, where u_{TV} is the solution of:

$$G_0(u) = \int_\Omega (\alpha|\nabla u| + \beta|u - u_0|^2)dxdy,$$

and $1 = v_{\rho=0} = \frac{1}{1+4\alpha\cdot 0\cdot|\nabla u|}$ from the case $p = 1$ (v_ρ is associated with the functional G_ρ).

3.3. Generalization of Ambrosio-Tortorelli and Shah approximations in two variables. In this subsection we want only to mention that using the previous techniques, we can also generalize the approximations Ambrosio-Tortorelli and Shah in two variables, in order to ansure the existenceand uniqueness of minimizers (u_ρ, v_ρ and the convergence. But, experimentally, these non-convex approximations give very satisfactory results. The numerical algorithms are unconditionally stable, this being related to the fact that the function ϕ is positive and increasing (note that also the Mumford-Shah functional does not satisfy the uniqueness of the minimizer).

Let ϕ_ρ be an arbitrary function satisfying the assumptions of Theorem 3.1. Then we have the existence of a convex function ψ_ρ, associated to ϕ_ρ by this theorem. Assume further that $\phi_\rho(t) \to t$, for all $t \geq 0$, as $\rho \to 0$. Then $\phi_\rho(t^p) \to t^p$, for all $t \geq 0$, as $\rho \to 0$, with for instance $p \geq 1$.

We can consider the following functional:

$$F_\rho(u, v) = \int_\Omega \Big(\alpha(v^2|\nabla u|^p + \psi_\rho(v)) + \beta|u - u_0|^2 + \rho|\nabla v|\Big)dxdy. \tag{3.6}$$

Note that we use the regularizing term $\rho|\nabla v|$, but which is now anisotropic, comparing with $\rho|\nabla v|^2$, in the Ambrosio-Tortorelli and Shah approximations, this being necessary to preserve better the edges in the image function u. Of-course, and as we already mentioned, this general approximation will be mathematically better well-posed than our reduced non-convex approximations.

We can also consider the following regularized functional in only one variable,but closely related to the previous-one:

$$F(u) = \int_\Omega \Big(\alpha\phi_\rho(|\nabla u|^p) + \rho|\nabla(\sqrt{\phi'(|\nabla u|^p)})| + \beta|u - u_0|^2\Big)dxdy. \tag{3.7}$$

The associated Euler equation will be of the fourth order. These last two functionals can be related to some higher order functionals proposed by Chambolle-Lions in [9].

The theoretical study of these two approximations, including the convergence to a given functional depending on the jump set S_u (the Γ-limit), in this general case, may be an interesting problem (the case $p = 1$ is related to [6] and of-course the cases with $p > 1$ will be related to the paper of Ambrosio-Tortorelli [3]).

4. Comparisons and numerical results. This section is devoted to comparisons and numerical results for the different approximations discussed before. We will not give the details of the numerical algorithms. We mention only that we use finite differences schemes, and the discretization of the divergence operators in each equation is based on the method of Rudin-Osher-Fatemi from [12]. The resulting algorithms, even for the non-convex cases, are unconditionally stable (we refer the reader to [4] and [15] for the details of the schemes). We recall the notations: $\phi_{\rho'}(t) = \frac{t}{1+4\alpha\rho' t} = \phi_\rho(t) = \frac{t}{1+\rho t}$ (changing $4\alpha\rho'$ in ρ for simplicity) and $\varphi(t) = \frac{1}{\rho}\log(1+\rho t)$.

We present in Fig. 4.1 our first synthetic test image (original and noisy). We compare the results using the SNR (the signal to noise ratio, between the original image and the result, after normalization of images). The SNR is small for a noisy image, and has to increase, for the reconstructed image.

We can first fix $\beta = 0.5$ in all models. Then, because there is not a rigorous choice of parameters, we use for each model the parameters for which we obtained the best results (the best SNR). However, we use the same parameters for the corresponding models: Ambrosio-Tortorelli model and the new model with ϕ and $p = 2$, also for the modified Shah model and the new model with ϕ and $p = 1$. For the initial Shah model and the new model using φ we have independent parameters. For the presented results, the number of iterations is also choosen in order to have the best SNR, in other words, we stop the iterative algorithms when we obtained the best SNR for the reconstructed image u. But, as we will see (for instance for the Ambrosio-Tortorelli approximation), v may be not smoothed enough for the corresponding and the best u. Using our new models, we will see that we can obtain very satisfactory results both for u and its associated $v(|\nabla u|)$, without performing more iterations only for v.

We present each time the reconstructed image u together with the SNR, and the corresponding or associated edge-detector v.

We show first the results obtained with the Ambrosio-Tortorelli, initial Shah and modified Shah models (Fig. 4.2). We mention that for the Ambrosio-Tortorelli model, we have to stop the algorithm very fast, because the image u becomes too much diffused, but in the same time v is not denoised enough. We remark that using the modified Shah model, the results are much better than with the Ambrosio-Tortorelli and initial Shah models. But the edges in v are still not so sharp.

Now, in Fig. 4.3 we show the results obtained with the new reduced models using the functions ϕ and φ. We remark that the new model still gives a better result (SNR=22.30 with ϕ, $p = 1$) than the modified Shah model (SNR=19.80). Moreover, the edges are better preserved with the new model. For this synthetic image, the two functions ϕ and φ produced similarly results. However, for the cases with $p > 1$, we obtained better results for the image function u using φ (the SNR is larger than using ϕ).

We have shown in a previous paper [8] that the locations of edges are very well preserved using the new model with ϕ, and also that the new algorithms are faster (we present in Fig. 4.4 the SNR function of the time).

In Fig. 4.5 and Fig. 4.6, we present the results using the new model with ϕ, $p = 1$: $(u, v(|\nabla u|))$, and the cross-sections for $y = 87$ respectively, for a decreasing sequence of ρ to 0, in order to illustrate the behavior of the solution as a function of ρ. We also add the result obtained for $\rho = 0$, this being exactly the total variation minimization. For this case, $v = 1$ everywhere. We remark that the parameter ρ has a regularizing effect on the solution u, but forces $v_\rho(|\nabla u|)$ to be different from $1 = \sqrt{\phi'_\rho(0)}$ only on the set

S_u. We see that there is an optimal value for the parameter ρ (here for $\rho = 0.0181$). For $0 \leq \rho < \rho_{optimal} = 0.0181$ here, the edges in u are too much regularized. This can be seen very well in the final result when $\rho = 0$ (TV). But these results illustrate also the behavior of v as $\rho \to 0$: we have experimentally that $v_\rho \to 1$ in the norm L^2, as $\rho \to 0$. We study more this behavior in the next two figures, where we present by graphs the SNR and MSE (mean square error) for "original-u_ρ" (Fig. 4.7), the MSE for "$u_{TV} - u_\rho$" and "$v_\rho - 1$" (Fig. 4.8). These show that the numerical solution u_ρ converges to $u_{\rho=0} = u_{TV}$ as $\rho \to 0$ and also the convergence of v_ρ to 1, as in the theoretical result of Ambrosio and Tortorelli.

We end the paper by showing results with an original noisy image representing an office. Our new reduced models still produce better results both for the reconstructed image u and the edge-detector v.

We remark that the results for u using the function φ are slightly better than using the first new model with ϕ, for $p > 1$. This is due to the different behavior of these functions to infinity. It will be interesting to compare these results with the proposed generalized and regularized approximations from (3.6) and (3.7).

REFERENCES

[1] L. AMBROSIO, *A compactness theorem for a special class of functions of bounded variation*, Boll. Un. Mat. Ital. (7), 3-B (1989), pp. 857-881.

[2] L. AMBROSIO AND V. TORTORELLI, *Approximation of Functionals Depending on Jumps by Elliptic Functionals via* $\Gamma-$*Convergence*, Comm. on Pure and Applied Mathematics, Vol. XLIII(1990), pp. 999-1036.

[3] L. AMBROSIO AND V. TORTORELLI, *On the Approximation of Free Discontinuity Problems*, Bollettino U.M.I. (7) 6-B (1992), 105-123.

[4] G. AUBERT AND L. VESE, *A variational method in image recovery*, SIAM J. Num. Analysis, Vol. 34, No. 5 October 1997, pp. 1948-1979.

[5] L. BLANC-FÉRAUD, S. TEBOUL, G. AUBERT AND M. BARLAUD, *Nonlinear regularization using constrained edges in image reconstruction*, IEEE Proc. of the 3rd International Conference of Image Processing, Lausanne 1996, Vol. II, pp.449-452.

[6] G. BOUCHITTÉ, C. DUBS, P. SEPPECHER, *Transitions de phases avec un potentiel dégénéré à l'infini, application á l'équilibre de petites gouttes*, C.R.A.S., Série I, Tome 323, No 9, Nov. 1996, pp.1103-1108.

[7] B. BOURDIN, *Image Segmentation With a Finite Element Method*, L.P.M.T.M. Report, Institut Galilée, Université de Paris Nord, 1997.

[8] T. CHAN AND L. VESE, *Variational Image Restoration & Segmentation Models and Approximations*, UCLA CAM Report 47, October 1997, submitted to IEEE Transactions in Image Processing.

[9] . A. CHAMBOLLE, P.L. LIONS, *Image recovery via total variation minimization and related problems*, Numer. Math. (1997) 76: pp 167-188.

[10] J.-M. MOREL AND S. SOLIMINI, *Variational Methods in Image Segmentation*, Progress in Nonlinear Differential Equations and Their Applications, Vol. 14, Birkhäuser Boston, 1995.

[11] D. MUMFORD AND J. SHAH, *Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems*, Communications on Pure and Applied Mathematics, Vol. XLVII(1989), pp. 577-685.

[12] L. RUDIN, S. OSHER, E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Physica D, 60(1992), pp. 259-268.

[13] J. SHAH, *A Common Framework for Curve Evolution, Segmentation and Anisotropic Diffusion*, Proceedings of I.E.E.E. Conference on Computer vision and pattern recognition, San Francisco, June 18-20, 1996, pp. 136-142.

[14] S. TEBOUL, L. BLANC-FÉRAUD, G. AUBERT AND M. BARLAUD, *Variational Approach for Edge-Preserving Regularization using Coupled PDE's*, submitted to IEEE Trans. on Image Processing, 1996.

[15] L. VESE, *Variational problems and PDE's for image analysis and curve evolution*, P.H.D. Thesis, University of Nice, Nov. 1996.

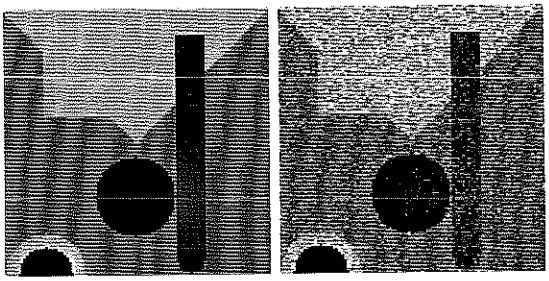

FIG. 4.1. *Original image and a noisy version* u_0 *(SNR(orig-noisy)=7.38db).*

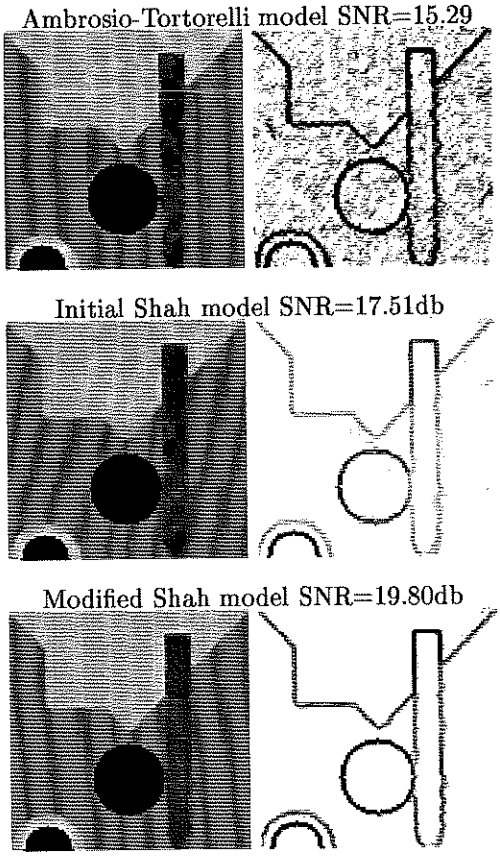

FIG. 4.2. *Ambrosio-Tortorelli, Initial Shah model, and modified Shah model, from top to bottom.*

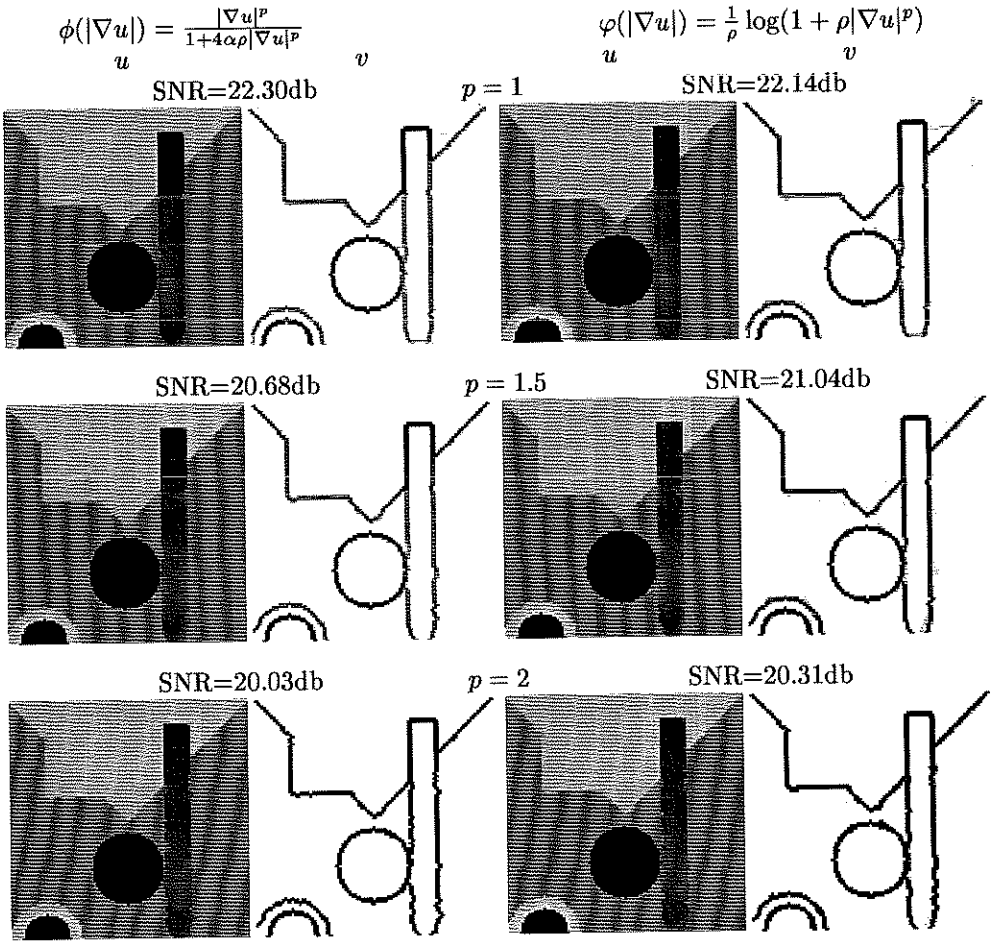

FIG. 4.3. *The new model using ϕ (left) and φ (right), for $p = 1;\ 1.5;\ 2$.*

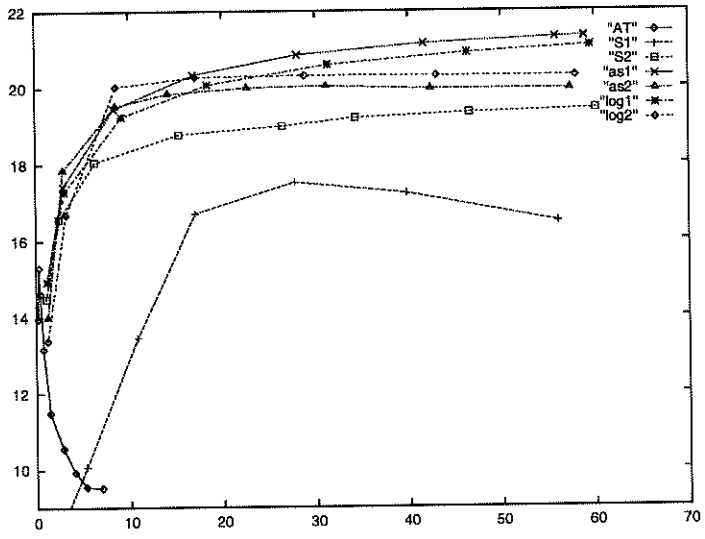

FIG. 4.4. *The SNR function of the time for the different models.*

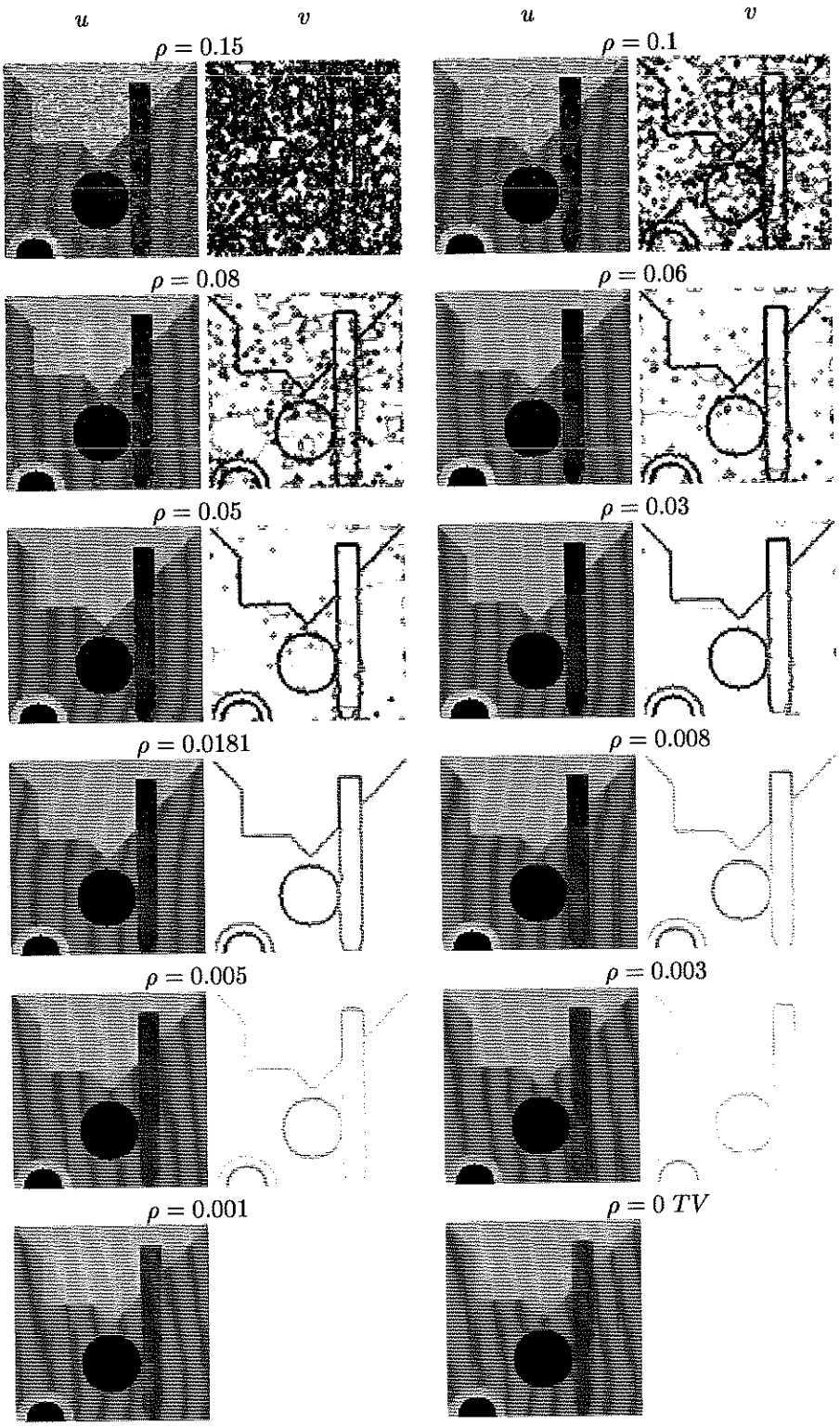

FIG. 4.5. *u and v as* $\rho \to 0$ *and* $\rho = 0$, *for* $p = 1$ *and* $\phi_\rho(t) = \frac{t}{1+\rho t}$.

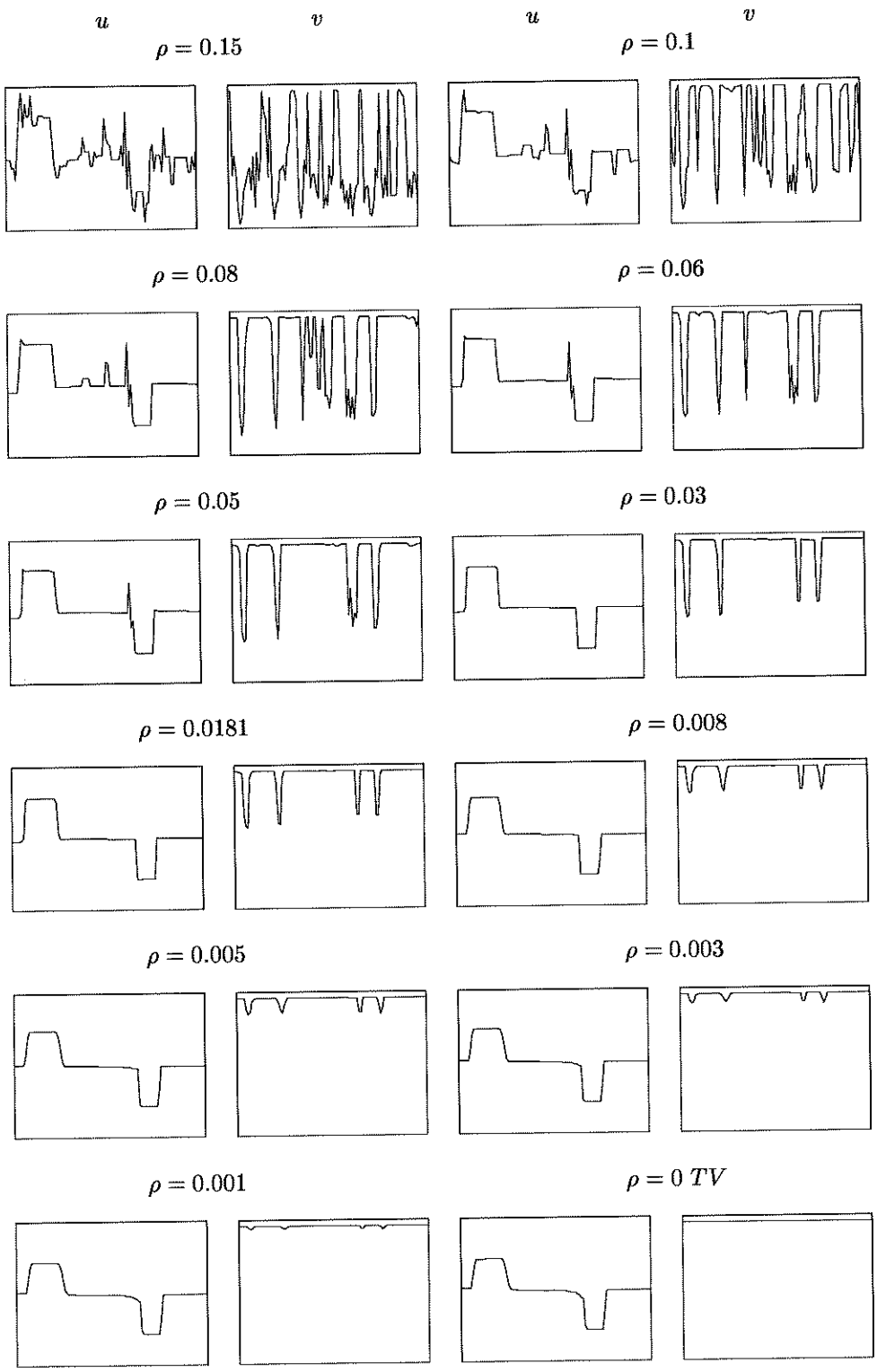

FIG. 4.6. *Associated cross-sections for* $y = 87$, *as* $\rho \to 0$ *and* $\rho = 0$.

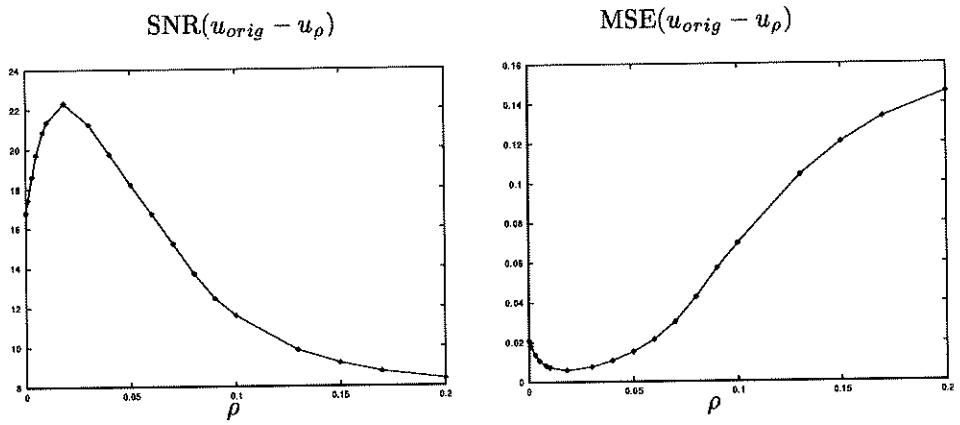

FIG. 4.7. *The SNR (left) and MSE (right) between the original image and the result u_ρ, function of ρ, as $\rho \to 0$ and for $\rho = 0$ (TV).*

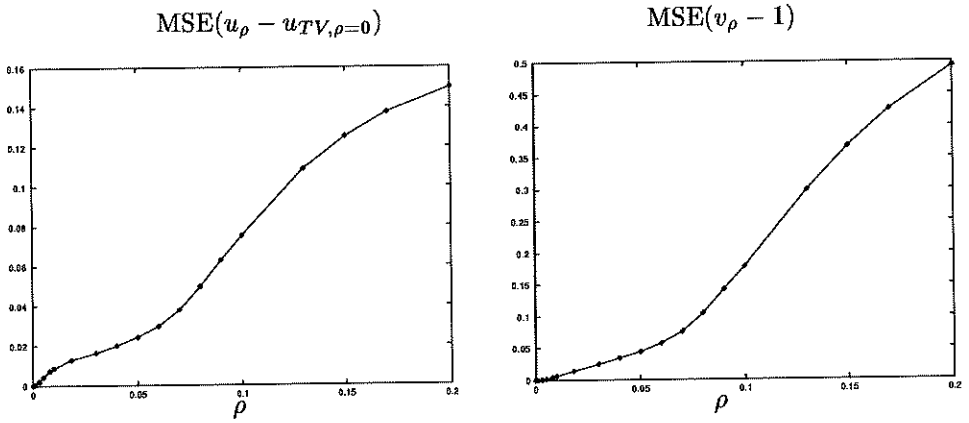

FIG. 4.8. *The MSE of $(u_\rho - u_{TV})$ (left) and of $(v_\rho - 1)$ (right) function of ρ, as $\rho \to 0$ and for $\rho = 0$ (TV).*

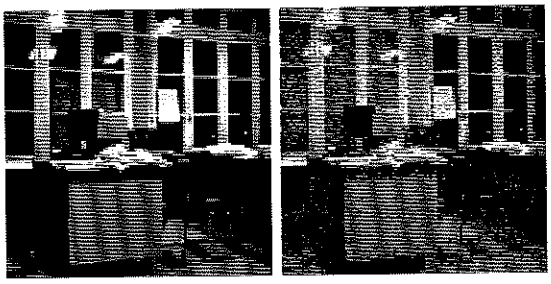

FIG. 4.9. *Real original image and noisy with SNR=11.08db.*

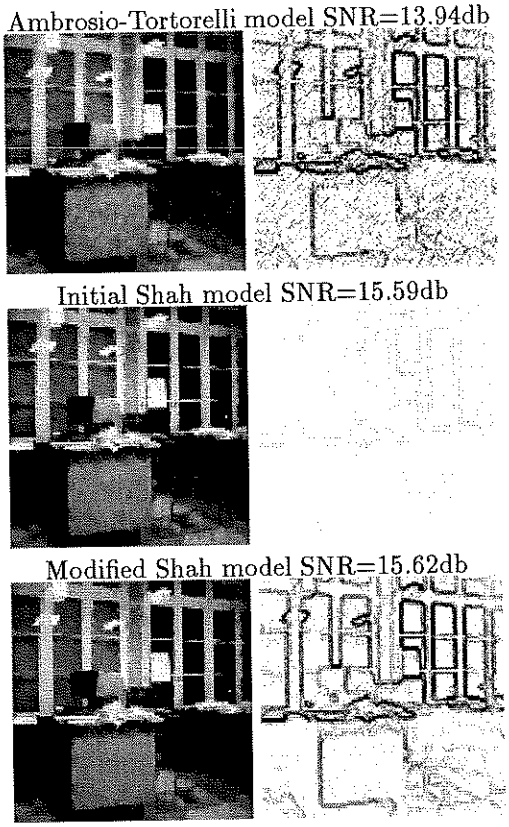

FIG. 4.10. *Results obtained with the Ambrosio-Tortorelli, Initial Shah and Modified Shah models, from top to bottom.*

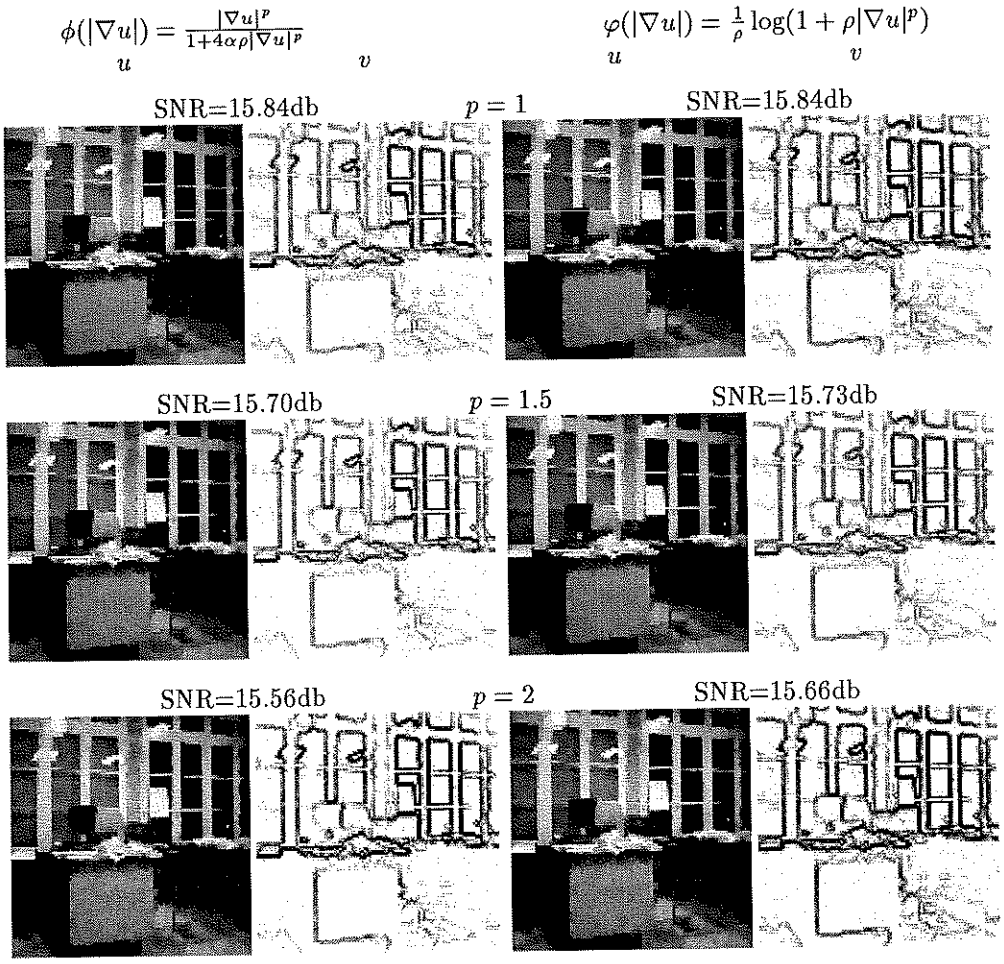

FIG. 4.11. *results obtained using the new reduced models with ϕ (left) and φ (right) for the real image.*