On the Control of a Double Obstacle Problem in Image Reconstruction

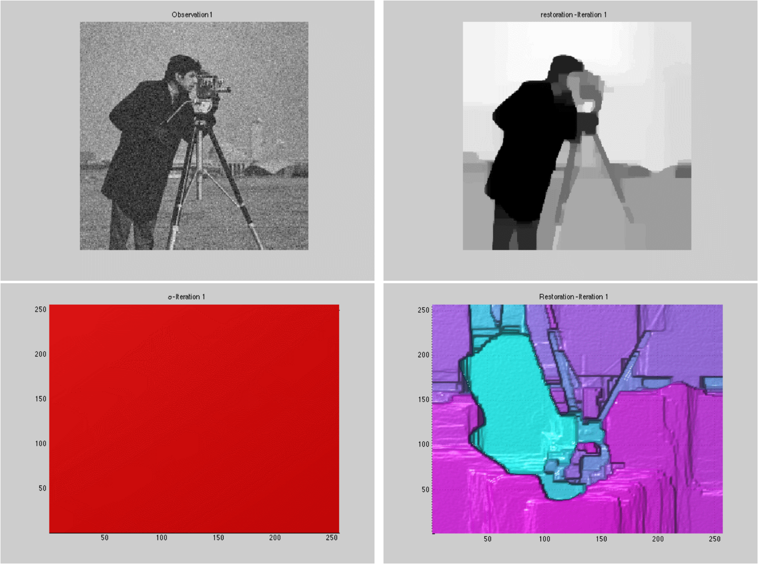
Carlos N. Rautenberg
(joint work with Michael Hintermüller, and Andreas Langer)

Humboldt-Universität zu Berlin

August 7, 2014

1. The Lower Level Poblem (TV predual)
2. The Bilevel Problem and Approximations
3. Numerical Tests
    - Triangle+Rectangle
    - Circle
    - Cameraman

Visual Incentive



The Lower Level Problem (TV regularization)

Given $f \in L^2(\Omega)$ where $f = u_{true} + \eta$, $\int_\Omega \eta = 0$ and $\int_\Omega |\eta|^2 = \sigma^2$. Consider $\alpha > 0$, the TV model reads:

$$\min_{u \in BV(\Omega)} \frac{1}{2} \int_\Omega |u - f|^2 + \alpha \int_\Omega |\mathcal{D}u|, \tag{TV}$$

where $\int_\Omega |\mathcal{D}u| := |\mathcal{D}u|(\Omega)$, the total mass of the Borel measure $\mathcal{D}u$ determined by the distributional gradient of u:

$$\int_\Omega |\mathcal{D}u| = \sup \left\{ \int_\Omega u \, \mathrm{div}\mathbf{v} \mathrm{d}x \,\middle|\, \mathbf{v} \in C_c^1(\Omega; \mathbb{R}^2),\ |\mathbf{v}(\mathbf{x})|_\infty \leq 1 \text{ a.e. } \mathbf{x} \in \Omega \right\}.$$

The solution to (TV) satisfies that for :

- α high, contains no noise but also details in u_{true} are lost.
- α small, details for u_{true} are retained but also (possibly) noise.

The spatially variant α

Given $f \in L^2(\Omega)$ and $\alpha : \Omega \to \mathbb{R}$, the TV model reads:

$$\min_{u \in BV(\Omega)} \frac{1}{2} \int_\Omega |u - f|^2 + \int_\Omega \alpha |\mathcal{D}u|. \tag{TV*}$$

- A proper choice of the spatially variant α could help recover small details in certain regions while also properly denoising flat regions .
- Well-posedness of the problem requires certain *regularity* of α: it should be $|\mathcal{D}u|$-measurable ($|\mathcal{D}u|$ is a Borel measure).
- Additionally, if α is not positive on $\overline{\Omega}$, the problem might be ill-posed.

The spatially variant α

Existence

If $\alpha \in C(\overline{\Omega})$ and $\alpha(x) > 0$ for all $x \in \overline{\Omega}$, then there is a unique solution to (TV*).

Therefore, the mapping

$$C^+(\overline{\Omega}) \ni \alpha \mapsto u_\alpha \in BV(\Omega),$$

is well-defined. However, we will look at u_α from the point of view of Fenchel duality for several reasons...

The (Fenchel) Pre-dual of (TV*)

Duality

Let $\alpha \in C(\overline{\Omega})$ and $\alpha(x) > 0$ for all $x \in \overline{\Omega}$. The Fenchel pre-dual problem of (TV*)

$$\min_{u \in BV(\Omega)} \frac{1}{2}\int_\Omega |u - f|^2 + \int_\Omega \alpha |\mathcal{D}u|,$$

is given by

$$\min_{\mathbf{p} \in H_0(\mathrm{div})} \frac{1}{2}|\mathrm{div}\mathbf{p} + f|^2_{L^2} \quad \text{s.t} \quad |\mathbf{p}(x)|_\infty \leq \alpha(x) \text{ a.e. } x \in \Omega, \quad (\mathrm{TV}^*_{pd})$$

and $u_\alpha = \mathrm{div}\mathbf{p}_\alpha + f$.

The result it is not a trivial extension of known results, it requires results based on density of closed, convex, sets...

A digression on the density of closed, convex sets

Let X be a space of $\mathbb{R}^M$-functions over $\Omega \subset \mathbb{R}^N$

$$\mathbb{K}(X) := \{\mathbf{f} \in X : |\mathbf{f}(x)| \leq \alpha(x) \text{ a.e., } x \in \Omega\}.$$

The previous theorem requires that $\overline{\mathbb{K}(\mathscr{D}(\Omega)^M)}^{H_0(\mathrm{div})} = \mathbb{K}(H_0(\mathrm{div}))$ and $\overline{\mathbb{K}(\mathscr{D}(\Omega)^M)}^{C_0(\Omega)^M} = \mathbb{K}(C_0(\Omega)^M)$.

This raises a general question: If X_0 is densely and continuously embeded on the Banach space X_1, is this sufficient to establish that

$$\overline{\mathbb{K}(X_0)}^{X_1} = \mathbb{K}(X_1)?$$

A digression on the density of closed, convex sets

Let X be a space of $\mathbb{R}^M$-functions over $\Omega \subset \mathbb{R}^N$

$$\mathbb{K}(X) := \{\mathbf{f} \in X : |\mathbf{f}(x)| \leq \alpha(x) \text{ a.e., } x \in \Omega\}.$$

The previous theorem requires that $\overline{\mathbb{K}(\mathscr{D}(\Omega)^M)}^{H_0(\mathrm{div})} = \mathbb{K}(H_0(\mathrm{div}))$ and $\overline{\mathbb{K}(\mathscr{D}(\Omega)^M)}^{C_0(\Omega)^M} = \mathbb{K}(C_0(\Omega)^M)$.

This raises a general question: If X_0 is densely and continuously embeded on the Banach space X_1, is this sufficient to establish that

$$\overline{\mathbb{K}(X_0)}^{X_1} = \mathbb{K}(X_1)?$$

The answer unfortunately is NO: in fact, you can find examples in which X_0 is continuously and densely embeded in $L^2(\Omega)$, but $\overline{\mathbb{K}(X_0)}^{L^2(\Omega)} = \{0\}$.

A digression on the density of closed, convex sets

Let X be a space of $\mathbb{R}^M$-functions over $\Omega \subset \mathbb{R}^N$

$$\mathbb{K}(X) := \{\mathbf{f} \in X : |\mathbf{f}(x)| \leq \alpha(x) \text{ a.e., } x \in \Omega\}.$$

The previous theorem requires that $\overline{\mathbb{K}(\mathscr{D}(\Omega)^M)}^{H_0(\text{div})} = \mathbb{K}(H_0(\text{div}))$ and $\overline{\mathbb{K}(\mathscr{D}(\Omega)^M)}^{C_0(\Omega)^M} = \mathbb{K}(C_0(\Omega)^M)$. ([Hintermüller, R.(2013)])

This raises a general question: If X_0 is densely and continuously embeded on the Banach space X_1, is this sufficient to establish that

$$\overline{\mathbb{K}(X_0)}^{X_1} = \mathbb{K}(X_1)?$$

The answer unfortunately is NO: in fact, you can find examples in which X_0 is continuously and densely embeded in $L^2(\Omega)$, but $\overline{\mathbb{K}(X_0)}^{L^2(\Omega)} = \{0\}$.

How choose α to get a good reconstruction?

Let $R : L^2(\Omega) \to L^\infty(\Omega)$ be defined [1] as

$$R(\mathrm{div}\mathbf{p})(x) := \int_\Omega w(x,y)\,(\mathrm{div}\mathbf{p})^2\,(y)\,\mathrm{d}y, \quad x \in \Omega,$$

with $\int_\Omega \int_\Omega w(x,y)\,\mathrm{d}y\,\mathrm{d}x = 1$ and $w(x,y) \geq 0$.

Let

$$x \mapsto M_1(\mathrm{div}\mathbf{p})(x) := \max(R(\mathrm{div}\mathbf{p}(x)) - \tilde{\sigma}^2, 0)^2,$$

and

$$x \mapsto M_2(\mathrm{div}\mathbf{p})(x) := \min(R(\mathrm{div}\mathbf{p}(x)) - \hat{\sigma}^2, 0)^2,$$

with some $\tilde{\sigma} = \sigma + \epsilon$ and $\hat{\sigma} = \sigma - \epsilon$.

[1] See ([Dong, Hintermüller, Rincón(2011)])

The Bilevel Problem

The problem of interest is then

$$\begin{aligned} &\text{minimize} \quad J(\alpha, \mathrm{div}\mathbf{p}) \\ &\text{over } (\mathbf{p}, \alpha) \in H_0(\mathrm{div}) \times C^+(\overline{\Omega}) \\ &\text{s.t.} \quad \alpha \in \mathcal{A}_{ad} \quad \text{and} \quad \mathbf{p} \text{ solving } (\mathrm{TV}^*_{pd}). \end{aligned}$$

where $\mathcal{A}_{ad} \subset C^+(\overline{\Omega})$ and

$$J(\alpha, \mathrm{div}\mathbf{p}) = \int_\Omega S_1(\alpha) M_1(\mathrm{div}\mathbf{p}) + \int_\Omega S_2(\alpha) M_2(\mathrm{div}\mathbf{p}),$$

where S_1 and S_2 are for scaling purposes.

Although existence of a solution might be obtained (using pre-compactness properties of $\mathcal{A}_{ad}$), algorithms to approximate solutions seem extremely hard to develop...

The Bilevel Problem

The map $\alpha \mapsto \mathbf{p}(\alpha)$ is complicated...

- Is $\mathcal{A}_{ad} \ni \alpha \mapsto \text{div}\mathbf{p}(\alpha)$ Lipschitz? It can be proven to be Lipschitz if $\mathcal{A}_{ad}$ comprises only "almost constant" functions...
- Is $\mathcal{A}_{ad} \ni \alpha \mapsto \mathbf{p}(\alpha)$ differentiable? ...
- Is $\mathbb{K} := \{\mathbf{q} \in H_0(\text{div}) : |\mathbf{q}(x)|_\infty \leq \alpha(x) \; a.e.\}$ polyhedric? If the control was in the forcing term of the problem, the differentiability question above is translated into the differentiability of the projection $\mathbf{q} \mapsto P_{\mathbb{K}}(\mathbf{q})$...

The Regularized Bilevel Problem

The problem of interest is then

$$\begin{aligned}&\text{minimize} \quad \mathcal{J}(\alpha, \operatorname{div}\mathbf{p}) := J(\alpha, \operatorname{div}\mathbf{p}) + \frac{\lambda}{2}|\alpha|^2_{H^1}\\ &\text{over } (\mathbf{p}, \alpha) \in H_0(\operatorname{div}) \times H^1(\Omega)\\ &\text{s.t.} \quad \alpha \in \mathcal{A}_{ad} \quad \text{and} \quad \mathbf{p} \text{ solving} \quad (\tilde{TV}^*_{pd}),\end{aligned}$$

where

$$\mathcal{A}_{ad} := \{\alpha \in H^1(\Omega) : 0 < \underline{\alpha} \le \alpha \le \overline{\alpha} < +\infty, \quad \text{a.e.}\},$$

and

$$\min_{\mathbf{p} \in H^1_0(\Omega)^l} \frac{\beta}{2}|\mathbf{p}|^2_{H^1_0} + \frac{1}{2}|\operatorname{div}\mathbf{p} + f|^2_{L^2} + \frac{1}{\epsilon}\mathfrak{P}(\mathbf{p}, \alpha). \qquad (\tilde{TV}^*_{pd})$$

The Regularized Bilevel Problem

- Nice First Order System (see slides of S. Ulbrich)
- The behaviour of the system as $(\beta, \epsilon) \downarrow (0, 0)$ may lead to something not useful at all.
- The solution mapping $H^1(\Omega) \ni \alpha \mapsto \mathbf{p}(\alpha) \in H_0^1(\Omega)^l$ of $(\tilde{TV}^*_{pd})$ is differentiable. It follows that the reduced objective map $\mathcal{F}(\alpha) := \mathcal{J}(\alpha, \text{div}\mathbf{p}(\alpha))$ is differentiable.

Projected Gradient + Armijo rule

Let $\alpha_0 \in \mathcal{A}_{ad}$ be in $H^2(\Omega) \cap C(\overline{\Omega})$ with $\tau \frac{\partial \alpha_0}{\partial \boldsymbol{\nu}} = 0$. Define $\{\alpha_k\}$ as

$$\alpha_{k+1} = P_{\mathcal{A}_{ad}}(\alpha_k - \tau_k \nabla \mathcal{F}(\alpha_k)), \quad k = 0, 1, \dots$$

where

- $P_{\mathcal{A}_{ad}} : H^1(\Omega) \to \mathcal{A}_{ad}$ is the minimum distance projection operator in the H^1-norm onto the closed convex set $\mathcal{A}_{ad}$.
- $\nabla \mathcal{F}(\alpha)$ denotes the gradient of $\mathcal{F}$ at $\alpha \in H^1(\Omega)$.
- $\{\tau_k\}$ is chosen according to (the general) Armijo's rule ([Bertsekas,Gafni(1982)]).

Preservation of Regularity

Preserved Regularity

Let $\Omega \subset \mathbb{R}^l$, $l = 1, 2$, be a bounded convex subset (or a polyhedron if $l = 3$) with $\underline{\alpha} < \overline{\alpha}$ regular enough and $\tau \frac{\partial \underline{\alpha}}{\partial \boldsymbol{\nu}} = \tau \frac{\partial \overline{\alpha}}{\partial \boldsymbol{\nu}} = 0$, where

$$\mathcal{A}_{ad} = \{\alpha \in H^1(\Omega) : 0 < \underline{\alpha} \leq \alpha \leq \overline{\alpha} \quad \text{a.e.}\}.$$

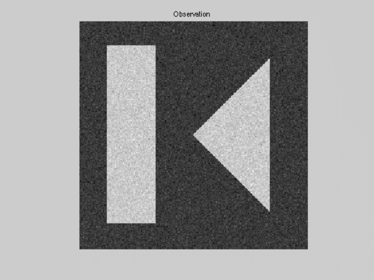
Then, the sequence $\{\alpha_k\}$ in $\mathcal{A}_{ad}$ generated by the Projected Gradient method preserves the initial iterate regularity:

$$\alpha_k \in H^2(\Omega) \cap C(\overline{\Omega}), \qquad k = 1, 2, \ldots.$$

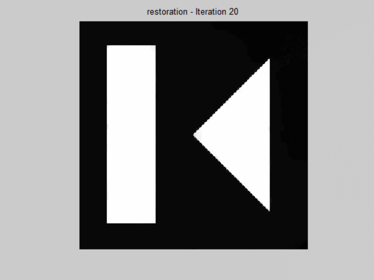
Furthermore, if $(\alpha^*, \mathbf{p}^*)$ is a solution to the regularized Bilevel problem, also $\alpha^* \in \mathcal{A}_{ad} \cap (H^2(\Omega) \cap C(\overline{\Omega}))$.

The convergence of $\{\alpha_k\}$ to a stationary point comes for free in [Bertsekas,Gafni(1982)].

The Triangle+Rectangle

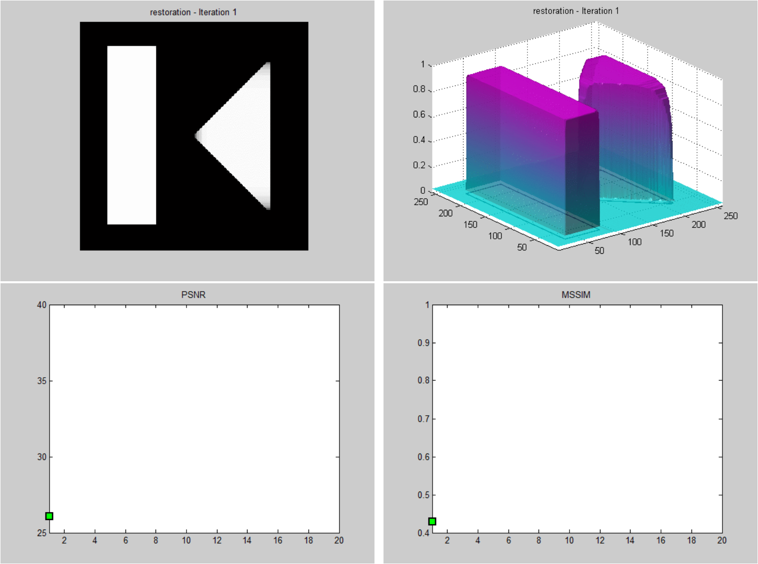


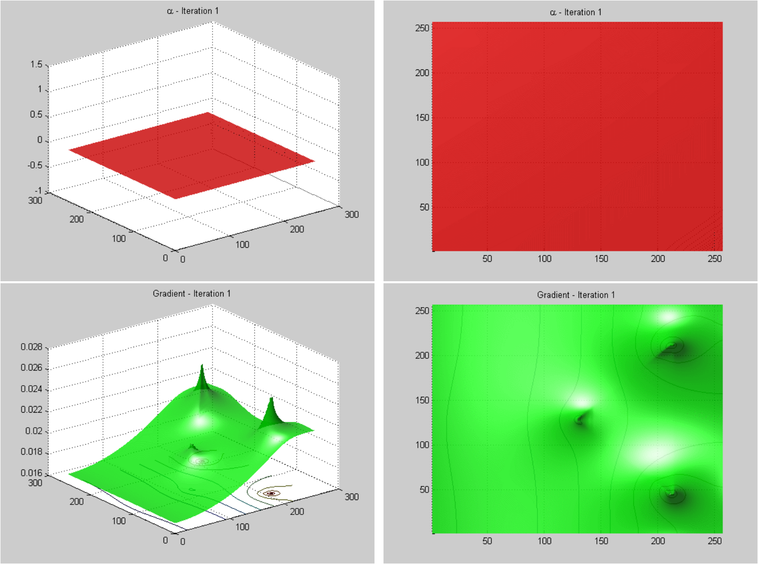
(a)



(b)

Figure : Noisy circle in (a) and restored circle (20 iterations) in (b)


restoration - Iteration 1
restoration - Iteration 1
1
0.8
0.6
0.4
0.2
0
250
200
150
100
50
50
100
150
200
250
PSNR
40
35
30
25
2
4
6
8
10
12
14
16
18
20
MSSIM
1
0.9
0.8
0.7
0.6
0.5
0.4
2
4
6
8
10
12
14
16
18
20


α - Iteration 1
1.5
1
0.5
0
-0.5
-1
300
200
100
0
0
100
200
300
α - Iteration 1
250
200
150
100
50
50
100
150
200
250
Gradient - Iteration 1
0.028
0.026
0.024
0.022
0.02
0.018
0.016
300
200
100
0
0
100
200
300
Gradient - Iteration 1
250
200
150
100
50
50
100
150
200
250

The Circle

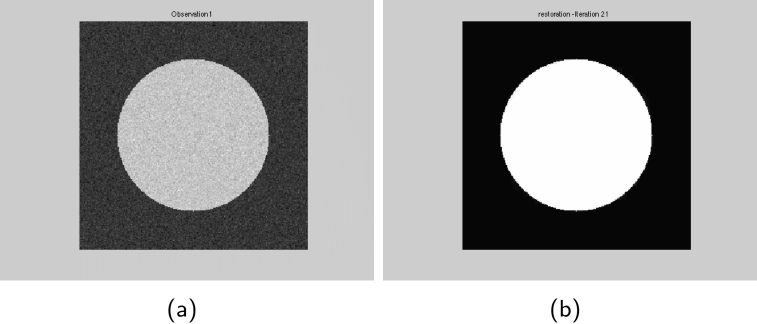
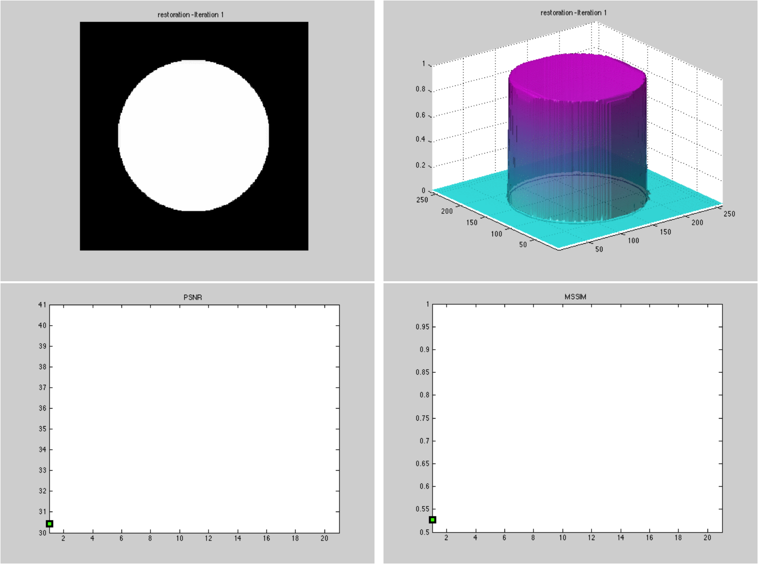
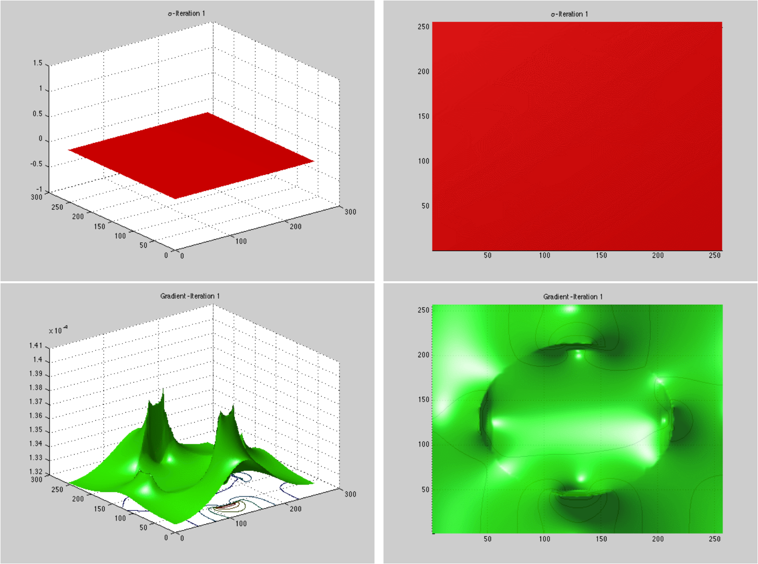


Figure : Noisy circle in (a) and restored circle (21 iterations) in (b)


restoration - Iteration 1
restoration - Iteration 1
PSNR
MSSIM


σ-Iteration 1
σ-Iteration 1
Gradient -Iteration 1
Gradient -Iteration 1

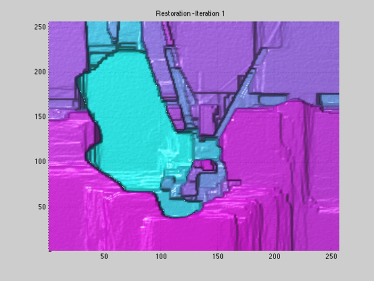
The Cameraman

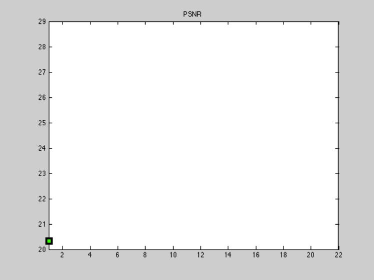
(a)

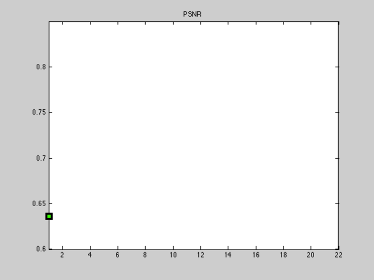
(b)

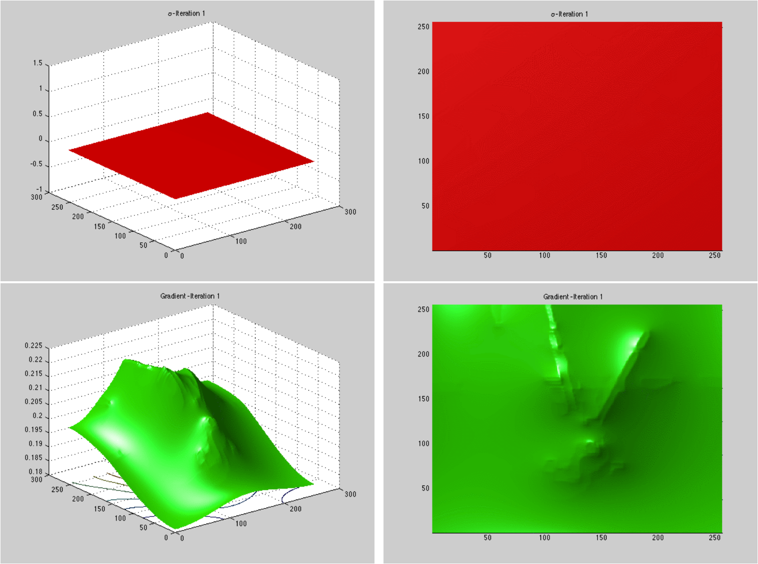
Figure : Noisy cameraman in (a) and restored cameraman (22 iterations) in (b)


restoration - Iteration 1


Restoration - Iteration 1
250
200
150
100
50
50
100
150
200
250


PSNR
29
28
27
26
25
24
23
22
21
20
2 4 6 8 10 12 14 16 18 20 22


PSNR
0.8
0.75
0.7
0.65
0.6
2 4 6 8 10 12 14 16 18 20 22


σ-Iteration 1
σ-Iteration 1
Gradient -Iteration 1
Gradient -Iteration 1

THANK YOU!