Improved ADMM based TV minimized Image Deblurring Without Boundary Artifacts

Dániel Hadházi, Áron Horváth, Gábor Horváth

Department of Measurement and Information Systems

Budapest University of Technology and Economics

MŰEGYETEM 1782

Contents of the presentation

- Basics of the image deblurring
  - Motivation of the TV minimization
- Introduce the proposed method
- Conclusions
  - Take some interesting conclusions of the deblurring
  - With visual examples

Image deblurring

- Main goal: invert degradation of acquisition
  - Linear shift invariant system + additive noise
  - PSF of the system and type of the noise are known
- Can be considered as a MAP estimation
  - $\arg\min_{\mathbf{x}} \left\{ -\log\left(\Pr\left(\mathbf{x}|\mathbf{y}\right)\right)\right\} = \arg\min_{\mathbf{x}} \left\{ \varphi\left(\mathbf{x},\mathbf{y}\right) + \kappa\left(\mathbf{x}\right)\right\}$
  - $\varphi\left(\mathbf{x},\mathbf{y}\right)$: loss term – penalizing the inconsistency between the deblurred and the input image
  - $\kappa\left(\mathbf{x}\right)$: regularizer term – ensures stability and defines some constraints

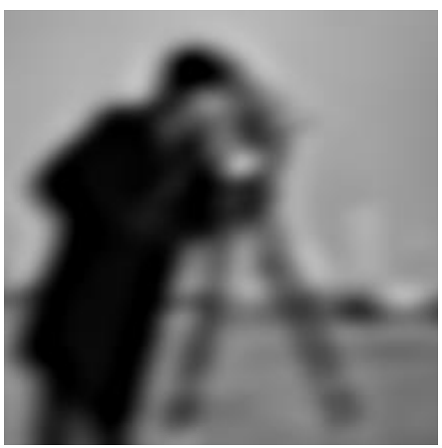
Image deblurring

- Loss term ($\varphi(\mathbf{x}, \mathbf{y})$):
  - Typically quadratic functions of the error
    - Assumption of additive observation Gaussian noise
  - Examined the Huber and the Sum of Absolute Error as well
    - With Gaussian observation noise
- Regularizer term ($\kappa(\mathbf{x})$):
  - PSF typically is a kind of low-pass filter
  - High noise sensitivity of the deblurring in high frequencies
  - Therefore the size of the gradient penalized with this term
  - Introduced domain constraints of the intensities

The size of the gradient

<u>Square of L2 norm:</u>

- Kind of ***Thikhonov*** regularization
- Similar to a low-pass post filtering
- The fine details of the picture are lost



<u>L2 norm:</u>

- ***Total Variation minimization*** based regularization
- Assumes that the gradient image is sparse
- Preserves edges and fine details

The proposed method

- The optimization is based on the Alternating Direction Method of Multipliers algorithm.
- Is an improved Split Bregman method:
  - Non negativity constraint is introduced on the intensities of the deblurred image.
  - The applied optimization method is modified in order to increase the rate of the convergence.
  - Defined a more general cost function that enables using different kind of loss functions.
  - Weighted loss terms are introduced that enables modelling different amount of noise for every pixel.

Formalization of the problem

- The formula of the primal optimization problem after applying the variable splitting:

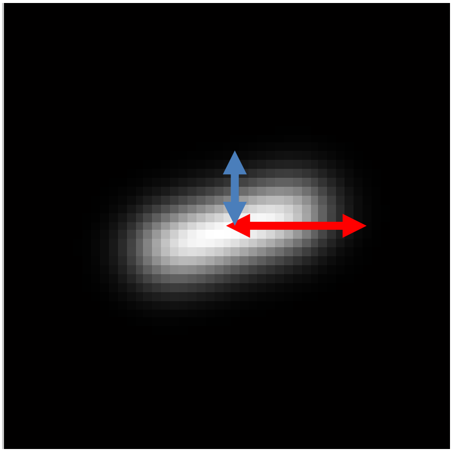
$$\min_{\mathbf{x},\mathbf{u},\mathbf{v},\mathbf{z},\mathbf{q}}. \quad \varphi_{\mathrm{w}}(\mathbf{u}) + \lambda \cdot \sum \left\| \left[\mathbf{v}_{(i,j)}, \mathbf{z}_{(i,j)} \right] \right\|_2 + I\left(\mathbf{q}_{(i,j)}\right)$$

$$\text{s.t.} \quad R\mathbf{x} = \mathbf{v},\ G\mathbf{x} = \mathbf{z};\ \mathbf{y} - H\mathbf{x} = \mathbf{u};\ \mathbf{q} = \mathbf{x}$$

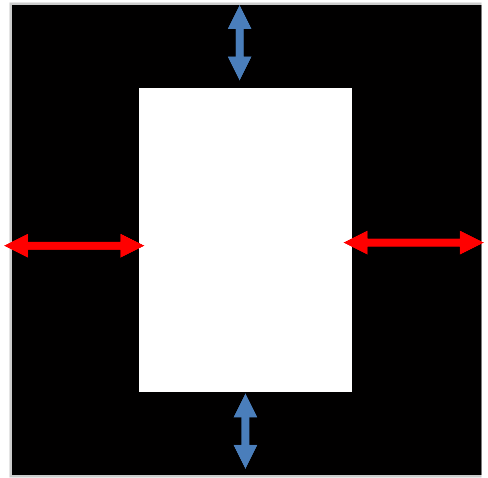
  - $\varphi_{\mathrm{w}}(\mathbf{u})$: is the loss function – penalizing the inconsistency
  - λ: is the weight term of the TV prior
  - $I(\cdot)$: is the domain constraint term – its value is infinite if the intensities of the deblurred picture are not nonnegative

Formalization of the problem

- The operators replaced by its circular counterparts:
  - The computations of the iterations can be done effectively
  - Cost of every iteration ~ 4× 2D FFT
  - If the weight matrix of the image error defined adequately than this change not introduce Boundary artifacts:



PSF

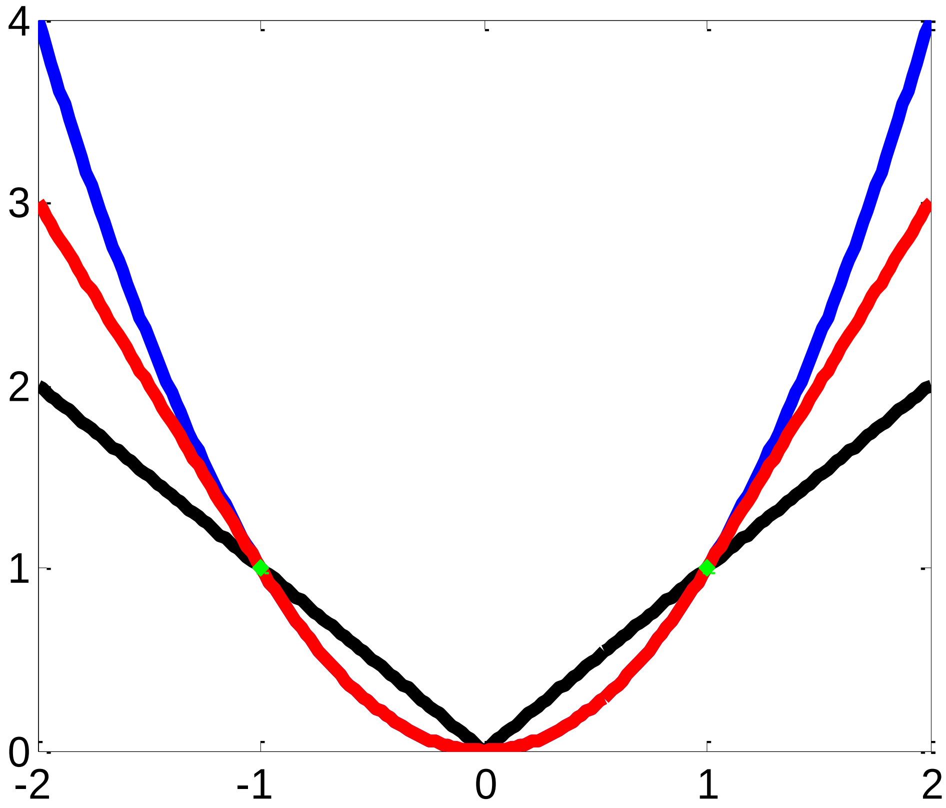


Weight matrix

Definition of the Loss term

- Weighted sum of square errors (SSE):
  - Zero mean Gaussian additive observation noise model
  - Moderate quality if there are high frequency textures
- Weighted sum of absolute errors (SAE):
  - Assumes that the additive noise is heavy-tailed
  - Theoretically corresponds to i.i.d Laplace noise
  - Better quantitative results, but high frequency artifacts
- Weighted Huber loss function (HLF):
  - Practically the mixture of the SAE and SSE

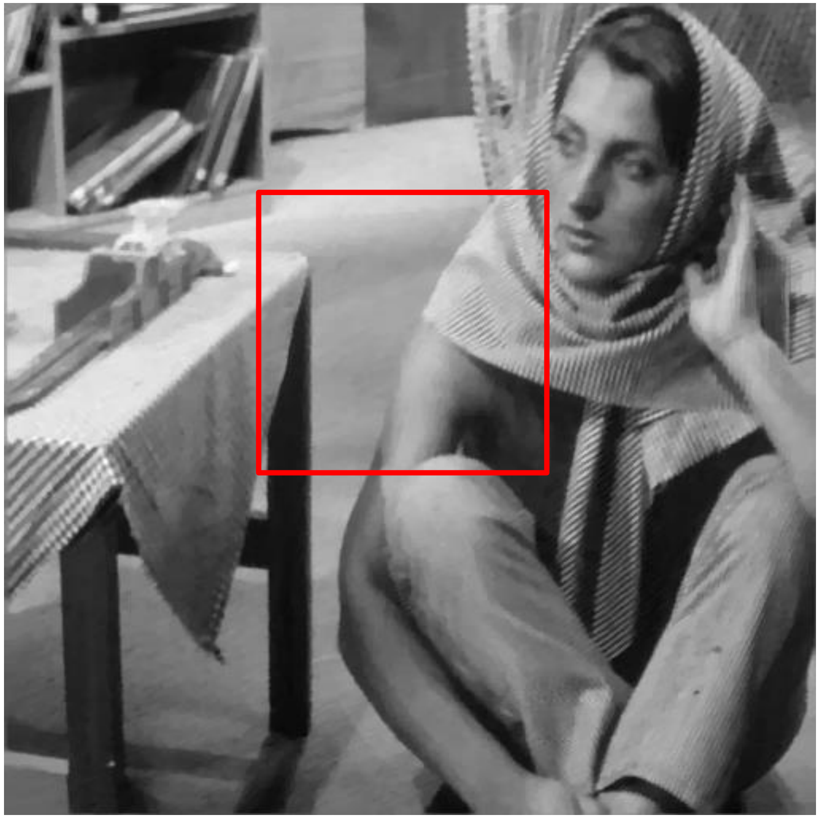
Definition of the Loss terms



- SSE
- SAE
- Huber
  - Change point of the definition is

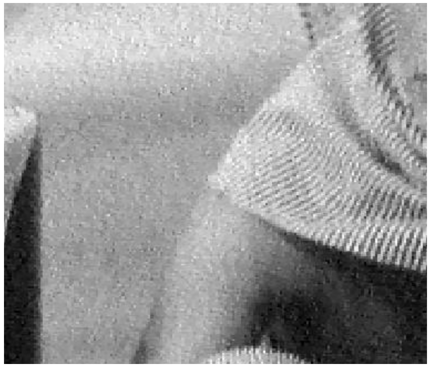
Quality as function of the loss term

- Let see an example – Gaussian noise is applied:



The deconvolved image

Sum of absolute errors

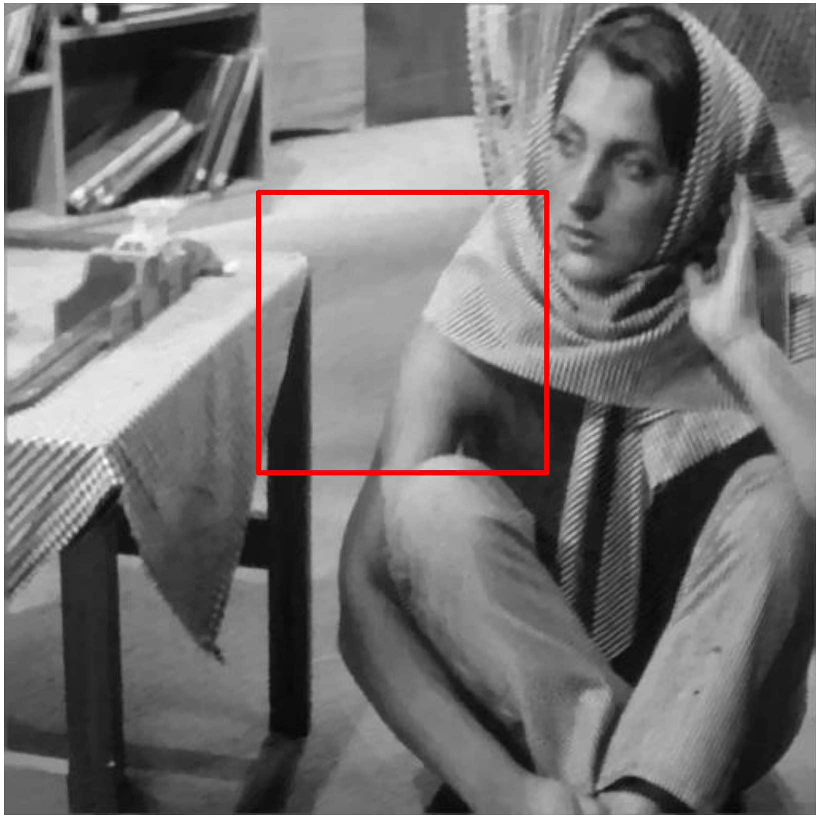


Huber penalty function

A Region of interest

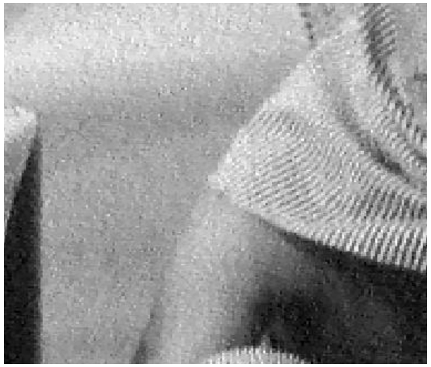
Quality as function of the loss term

- Let see an example – Gaussian noise is applied:



The deconvolved image

Sum of squared errors



Huber penalty function

A Region of interest

Conclusions of the experiments

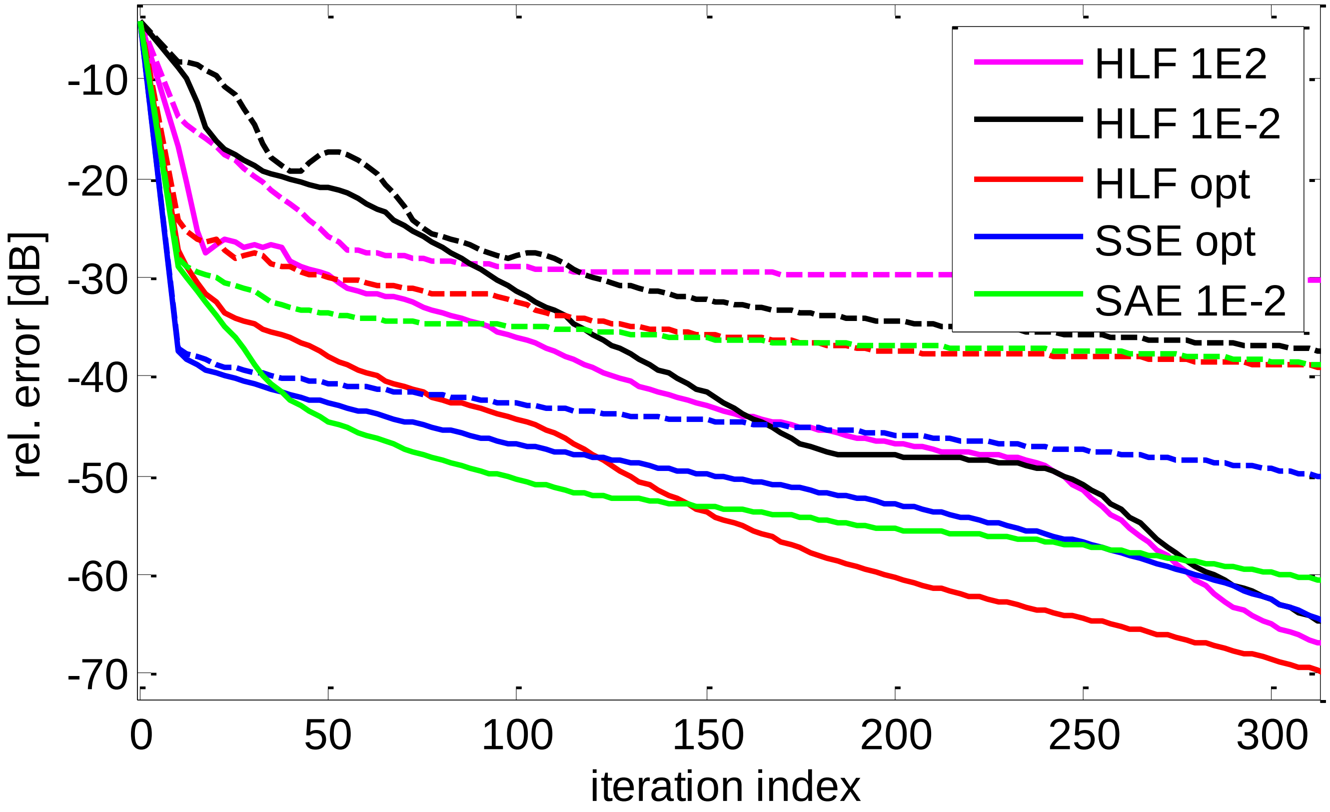
- Quantitative comparison with benchmark images:
  - Babara, Goldhill, Cameraman and Lena images
  - Additive Gaussian noise with different variance
  - Distortion: uniform blur / linear motion blur
  - ISNR was applied for quantifying the results
- General conclusions:
  - Better quality of images if the domain constraint used
  - SSE is only better than SAE in too noisy cases (BSNR<25 dB)
  - In every test cases there were a Huber function based solution which was the best
    - There exist change point value which …

Convergence properties of the optimization

- Convergence properties of the ADMM:
  - If the objective function is proper, closed and convex it converges to the solution in finite number of iterations
  - The speed significantly depends on the value of the hyper-parameters (AL penalty weights)
- How to calculate the optimal values:
  - Technically and theoretically it's very difficult (impossible?)
  - Instead of this, these weights are adjusted automatically at the end of every iteration:
    - In order to equalize the norms of the primal and the dual residuals

Automatic penalty weight modification

- Strictly accelerates the convergence rate in the case of every examined loss term:



Thank you for your kind attention!

Alternating Direction Method of Multipliers (ADMM)

- An improved iterative dual ascent method:
  - Utilize the idea of dual decomposition
  - Improvement by introducing new tags to the cost
    - Penalizing the primal feasibility gap
  - Effective if the criteria function can be decomposed into easily optimizable parts
    - E.g. parts which minimum can be calculated analytically
  - Also the optimization method of the Split Bregman algorithm

Alternating Direction Method of Multipliers (ADMM)

- Let see an example:
  - First step – introduce penalty of primal feasibility gap:

$$\begin{aligned} \min.\quad & f(\mathbf{x}) + g(\mathbf{y}) \\ \text{s.t.}\quad & \mathbf{y} = \mathbf{x} \end{aligned} \quad \Rightarrow \quad \begin{aligned} \min.\quad & f(\mathbf{x}) + g(\mathbf{y}) + (\rho/2)\cdot\|\mathbf{y}-\mathbf{x}\|_2^2 \\ \text{s.t.}\quad & \mathbf{y} = \mathbf{x} \end{aligned}$$

  - Then the dual problem optimized iteratively:

$$\max_{\boldsymbol{\eta}} \min_{\mathbf{x},\mathbf{y}} f(\mathbf{x}) + g(\mathbf{y}) + (\rho/2)\cdot\|\mathbf{y}-\mathbf{x}+\boldsymbol{\eta}\|_2^2$$

1. Minimizing over the primal variables one by one
2. Updating the dual variable by gradient ascend step
3. Optional – adjustment of the AL penalty weight