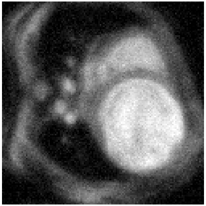
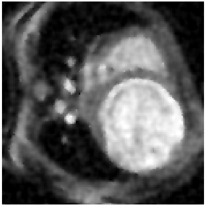
Constrained TV-based regularization framework

Camille Couprie*, Hugues Talbot*, Jean-Christophe Pesquet*
Laurent Najman* and Leo Grady†

* Université Paris-Est
Laboratoire d'Informatique Gaspard Monge
† Siemens Corporate Research, Princeton

ICASSP, 28 May 2011

Motivation


$\longrightarrow$


- Total Variation (TV) minimization : good regularization tool
- Weighted TV : penalization of the gradient leading to improved results

Our contribution

- General combinatorial formulation of the dual TV problem : easily suitable to various graphs
- Generic constraint in the dual problem : more flexible penalization of the gradient $\rightarrow$ sharper results

Outline

1. Generalization of TV models
2. Parallel Proximal Algorithm as an efficient solver
3. Results

Total variation regularization for image denoising

- Given an original image f
- Deduce a restored image u

Weighted anisotropic TV model [Gilboa and Osher 2007]

$$\min_u \underbrace{\int \left(\int w_{x,y}(u_y - u_x)^2 dy \right)^{1/2} dx}_{\text{regularization} \quad R(u)} + \underbrace{\frac{1}{2\lambda} \int (u_x - f_x)^2 dx}_{\text{data fidelity} \quad D(u)}$$

where

- $\lambda \in]0, +\infty[$ regularization parameter

Equivalent dual formulation

Weighted anisotropic TV model [Gilboa and Osher 2007]

$$\min_u \int \left(\int w_{x,y}(u_y - u_x)^2 dy \right)^{1/2} dx + D(u)$$

is equivalent [Chan, Golub, Mulet 1999] to the min-max problem

$$\min_u \max_{||p||_\infty \leq 1} \int \int w_{x,y}^{1/2}(u_y - u_x) p_{x,y} dxdy + D(u)$$

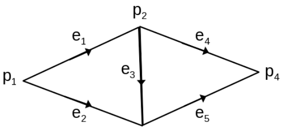
with p a projection vector field.

Main idea

- p was introduced in practice to compute a faster solution
- constraining p can promote better results

Discrete formulation on graphs - notations

Graph of N vertices, M edges



Incidence matrix $A \in \mathbb{R}^{M \times N}$

$$A = \begin{array}{c|cccc|} & p_1 & p_2 & p_3 & p_4 \\ \hline e_1 & -1 & 1 & 0 & 0 \\ e_2 & -1 & 0 & 1 & 0 \\ e_3 & 0 & -1 & 1 & 0 \\ e_4 & 0 & -1 & 0 & 1 \\ e_4 & 0 & 0 & -1 & 1 \end{array}$$

- A gradient operator
- $A^\top$ divergence operator
- allows general formulation of problems on arbitrary graphs

For more details : L. Grady and J.R. Polimeni,

"Discrete Calculus : Applied Analysis on Graphs for Computational Science", Springer, 2010.

Discrete formulations of TV and its dual

Let $u \in \mathbb{R}^N$ be the restored image.
[Bougleux *et al.* 2007]

$$\min_u \sum_{i=1}^{n} \Big(\sum_{j \in N_i} w_{i,j}(u_j - u_i)^2 \Big)^{1/2} + D(u)$$

where $N_i = \{j \in \{1, \ldots, n\} \mid e_{i,j} \in E\}$.

We introduce the following combinatorial formulation for the primal dual problem

$$\min_u \max_{\|p\|_\infty \leq 1,\ p \in \mathbb{R}^M} p^\top ((Au) \cdot \sqrt{w}) + D(u)$$

Dual constrained TV based formulation

Constraining the projection vector

- Introducing the projection vector $F \in \mathbb{R}^M = p \cdot \sqrt{w}$
- Constraining F to belong to a convex set C

$$\min_{u \in \mathbb{R}^N} \underbrace{\sup_{F \in C} \quad F^\top (Au)}_{\text{regularization}} + \underbrace{\frac{1}{2\lambda} \|u - f\|_2^2}_{\text{data fidelity}}$$

- $C = \cap_{i=1}^{m-1} C_i \neq \varnothing$ where $C_1, \ldots, C_{m-1}$ closed convex sets of $\mathbb{R}^M$.
- Given $g \in \mathbb{R}^N$, $\theta_i \in \mathbb{R}^M$, $\alpha \geq 1$, $C_i = \{F \in \mathbb{R}^M \mid \|\theta_i \cdot F\|_\alpha \leq g_i\}$.

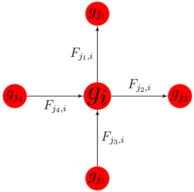
Dual constrained TV based formulation

$$\min_{u\in\mathbb{R}^N} \underbrace{\sup_{F\in C} \quad F^\top(Au)}_{\text{regularization}} + \underbrace{\frac{1}{2\lambda}\|u-f\|_2^2}_{\text{data fidelity}}$$

- $C = \cap_{i=1}^{m-1} C_i, \quad C_i = \{F \in \mathbb{R}^M \mid \|\theta_i \cdot F\|_\alpha \leq g_i\}, \ \alpha \geq 1.$

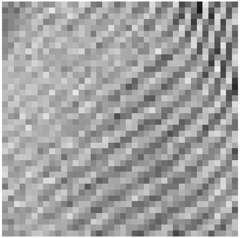
Example adapted to image denoising

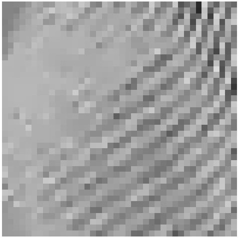
- $g_i \in \mathbb{R}^N$ weight on vertex i, inversely function of the gradient of f at node i.
- Flat area : weak gradient $\rightarrow$ strong g_i $\rightarrow$ strong $F_{i,j}$ $\rightarrow$ weak local variations of u.
- Contours : strong gradient $\rightarrow$ weak g_i $\rightarrow$ weak $F_{i,j}$ $\rightarrow$ large local variations of u allowed.

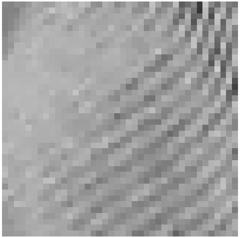


$$C_i = \{F \in \mathbb{R}^M \mid \sqrt{\sum_{j\in N_i} F_{j,i}^2} \leq g_i\}$$

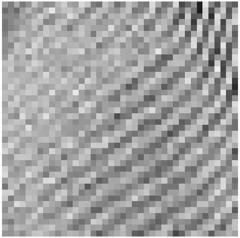
Sharper results

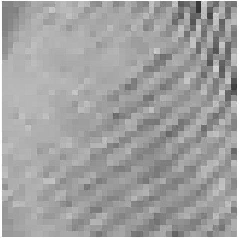

Noisy image

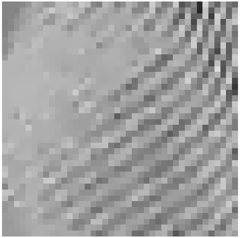

DCTV


Weighted TV

Sharper results


Noisy image


Weighted TV


DCTV

Extension of our DCTV based formulation

$$\min_{u \in \mathbb{R}^N} \underbrace{\sup_{F \in C} \quad F^\top (Au)}_{\text{regularization}} + \underbrace{\frac{1}{2\lambda} \|u - f\|_2^2}_{\text{data fidelity}}$$

- $f \in \mathbb{R}^Q$, observed image
- $u \in \mathbb{R}^N$, restored image
- $F \in \mathbb{R}^M$, dual solution : projection vector

Extension of our DCTV based formulation

$$\min_{u\in\mathbb{R}^N} \underbrace{\sup_{F\in C} \quad F^\top(Au)}_{\text{regularization}} + \underbrace{\frac{1}{2\lambda}\|Hu - f\|_2^2}_{\text{data fidelity}}$$

- $f \in \mathbb{R}^Q$, observed image
- $u \in \mathbb{R}^N$, restored image
- $F \in \mathbb{R}^M$, dual solution : projection vector
- $H \in \mathbb{R}^{Q\times N}$, degradation matrix

Extension of our DCTV based formulation

$$\min_{u\in\mathbb{R}^N} \underbrace{\sup_{F\in C} \quad F^\top(Au)}_{\text{regularization}} + \underbrace{\frac{1}{2\lambda}\|Hu - f\|_2^2 + \frac{\eta}{2}\|Ku\|^2}_{\text{data fidelity}}$$

- $f \in \mathbb{R}^Q$, observed image
- $u \in \mathbb{R}^N$, restored image
- $F \in \mathbb{R}^M$, dual solution : projection vector
- $H \in \mathbb{R}^{Q\times N}$, degradation matrix
- $K \in \mathbb{R}^{N\times N}$: projection onto Ker H , $\eta \geq 0$

Extension of our DCTV based formulation

$$\min_{u \in \mathbb{R}^N} \underbrace{\sup_{F \in C} \quad F^\top (Au)}_{\text{regularization}} + \underbrace{\frac{1}{2}(Hu - f)^\top \Lambda^{-1}(Hu - f) + \frac{\eta}{2}\|Ku\|^2}_{\text{data fidelity}}$$

- $f \in \mathbb{R}^Q$, observed image
- $u \in \mathbb{R}^N$, restored image
- $F \in \mathbb{R}^M$, dual solution : projection vector

- $H \in \mathbb{R}^{Q \times N}$, degradation matrix
- $K \in \mathbb{R}^{N \times N}$, projection onto Ker H , $\eta \geq 0$
- $\Lambda \in \mathbb{R}^{Q \times Q}$, matrix of weights, positive definite

Primal formulation

$$\min_{u\in\mathbb{R}^N} \quad \underbrace{\sigma_C(Au)}_{\text{regularization}} \quad + \underbrace{\frac{1}{2}(Hu-f)^\top\Lambda^{-1}(Hu-f)+\frac{\eta}{2}\|Ku\|^2}_{\text{data fidelity}}$$

- $C=\cap_{i=1}^{m-1}C_i\neq\varnothing$ where $C_1,\ldots,C_{m-1}$ closed convex sets of $\mathbb{R}^M$.
- σ_C support function of the convex set C

$$\sigma_C:\mathbb{R}^M\to]-\infty,+\infty]:a\mapsto\sup_{F\in C}F^\top a.$$

Dual problem

- The problem admits a unique solution $\widehat{u}$.
- Fenchel-Rockafellar dual problem :

$$\min_{F\in\mathbb{R}^M} \sum_{i=1}^{m-1} \underbrace{\iota_{C_i}(F)}_{f_i(F)} + f_m(F)$$

where ι_C is the indicator function of the convex C (equal to 0 inside C and $+\infty$ outside),
$f_m \colon F \mapsto \frac{1}{2} F^\top A \Gamma A^\top F - F^\top A \Gamma H^\top \Lambda^{-1} f$,
and $\Gamma = (H^\top \Lambda^{-1} H + \eta K)^{-1}$.

- If $\widehat{F}$ is a solution to the dual problem,

$$\widehat{u} = \Gamma \left(H^\top \Lambda^{-1} f - A^\top \widehat{F} \right).$$

Parallel ProXimal Algorithm (PPXA) optimizing DCTV

[Pesquet, Combettes, 2008]

$\gamma > 0, \nu \in]0,2[$.

Repeat until convergence

For (in parallel) $r = 1, \dots, s+1$

$$\pi_r = \begin{cases} \mathrm{P}_{C_r}(y_r) & \text{if } r \le s \\ (\gamma A \Gamma A^\top + I)^{-1}(\gamma A \Gamma H^\top \Lambda^{-1} f + y_{s+1}) & \text{otherwise} \end{cases}$$

$$z = \frac{2}{s+1}(\pi_1 + \cdots + \pi_{s+1}) - F$$

For (in parallel) $r = 1, \dots, s+1$

$$y_r = y_r + \nu(z - p_r)$$

$$F = F + \frac{\nu}{2}(z - F)$$

Parallel ProXimal Algorithm (PPXA) optimizing DCTV

[Pesquet, Combettes, 2008]

$\gamma > 0, \nu \in]0, 2[$.

Repeat until convergence

For (in parallel) $r = 1, \ldots, s+1$

$$\pi_r = \begin{cases} \mathrm{P}_{C_r}(y_r) & \text{if } r \le s \\ (\gamma A \Gamma A^\top + I)^{-1}(\gamma A \Gamma H^\top \Lambda^{-1} f + y_{s+1}) & \text{otherwise} \end{cases}$$

$z = \frac{2}{s+1}(\pi_1 + \cdots + \pi_{s+1}) - F$

For (in parallel) $r = 1, \ldots, s+1$

$y_r = y_r + \nu(z - p_r)$

$F = F + \frac{\nu}{2}(z - F)$

- Simple projections onto hyperspheres

Parallel ProXimal Algorithm (PPXA) optimizing DCTV

[Pesquet, Combettes, 2008]

$\gamma > 0, \nu \in]0,2[$.

Repeat until convergence

For (in parallel) $r = 1, \ldots, s+1$

$$\pi_r = \begin{cases} \mathrm{P}_{C_r}(y_r) & \text{if } r \leq s \\ (\gamma A \Gamma A^\top + I)^{-1}(\gamma A \Gamma H^\top \Lambda^{-1} f + y_{s+1}) & \text{otherwise} \end{cases}$$

$$z = \frac{2}{s+1}(\pi_1 + \cdots + \pi_{s+1}) - F$$

For (in parallel) $r = 1, \ldots, s+1$

$$y_r = y_r + \nu(z - p_r)$$

$$F = F + \frac{\nu}{2}(z - F)$$

- Linear system resolution

Quantitative perfomances

- Speed : competitive with the most efficient algorithm for optimizing weighted TV
- Denoising a 512×512 image
  - with an Alternated Direction of Multiplier Method : 0.4 seconds
  - with the Parallel Proximal Algorithm : 0.7 seconds
- Quantitative denoising experiments on standard images show improvements of SNR (from 0.2 to 0.5 dB) for images corrupted with Gaussian noise of variance σ^2 from 5 to 25.

Results in image denoising

Original image

Noisy SNR=10.1dB

Weighted TV SNR=13.4dB

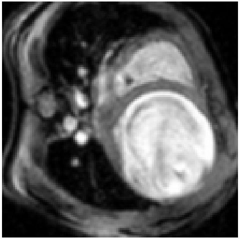
DCTV SNR=13.8dB

Results in image denoising

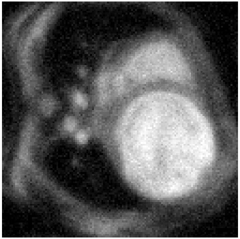
Weighted TV SNR=13.4dB

DCTV SNR=13.8dB

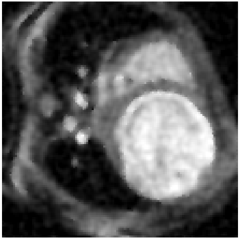
Image denoising and deconvolution



Original image



Noisy, blurred image SNR=12.3dB



DCTV result SNR=17.2dB

Image fusion

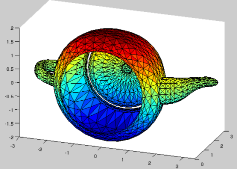
Original image

Noisy
SNR=7.2dB

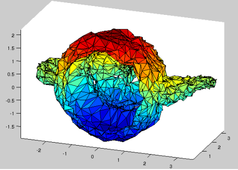
blurry
SNR=11.6dB

DCTV
SNR=16.3dB

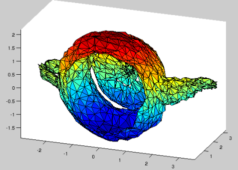
Mesh denoising



Original mesh

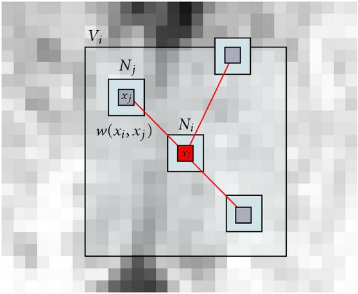


Noisy mesh



DCTV regularization on spatial coordinates

Non-local regularization



Non-local graph

Figure from P. Coupé et al.

Original image

Noisy PSNR=28.1dB

Nonlocal DCTV PSNR=35 dB

Conclusion

- Extension of TV models by generalization of the constraint on projection variable in the dual formulation
- Improved results
- Proposed algorithm efficiently solves convex problems involving the support function of an intersection of convex sets
- Application to arbitrary graphs