Proximal splitting methods for depth estimation.

Mireille EL GHECHE

Joint work with J.-C. Pesquet[1], J. Farah[2], Caroline Chaux[1]
and Béatrice Pesquet-Popescu[3]

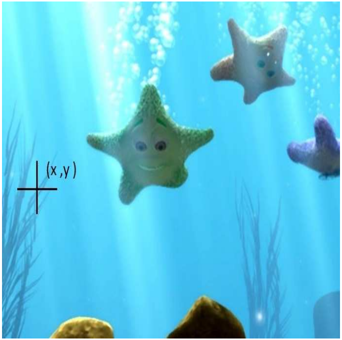
[1]Laboratoire d'Informatique Gaspard Monge - UMR CNRS 8049, Université Paris-Est, France.
[2]Departement of Telecommunications, Faculty of Engineering, Holy-Spirit University of Kaslik, Lebanon.
[3]Département Traitement du Signal et des Images, Telecom-ParisTech, France.
[4]Ecole doctorale science et technologie, Université Libanaise, Centre AZM, Tripoli, Lebanon.

"Journée des doctorants", 12 June 2012

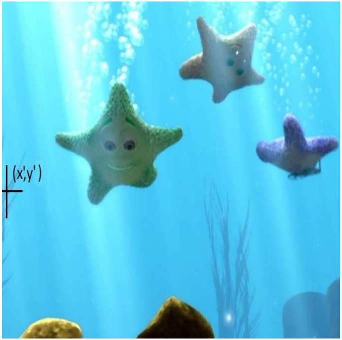
Outline

- Stereoscopic basics
- Energy model
  - Problem statement
  - Set theoretic estimation
  - Convex constraints
  - Sub-gradient projection method
- Proximal method
  - Proximity operator
  - PPXA+ algorithm
- Results

Disparity



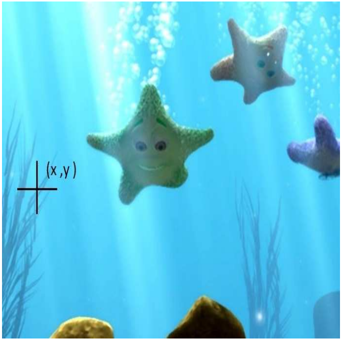
Left image (I_L)



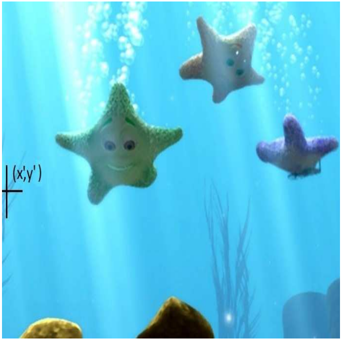
Right image (I_R)

$$u(x,y) = (x - x', y - y') \underset{y=y'}{\Longrightarrow} u(x,y) = (x - x')$$

Disparity



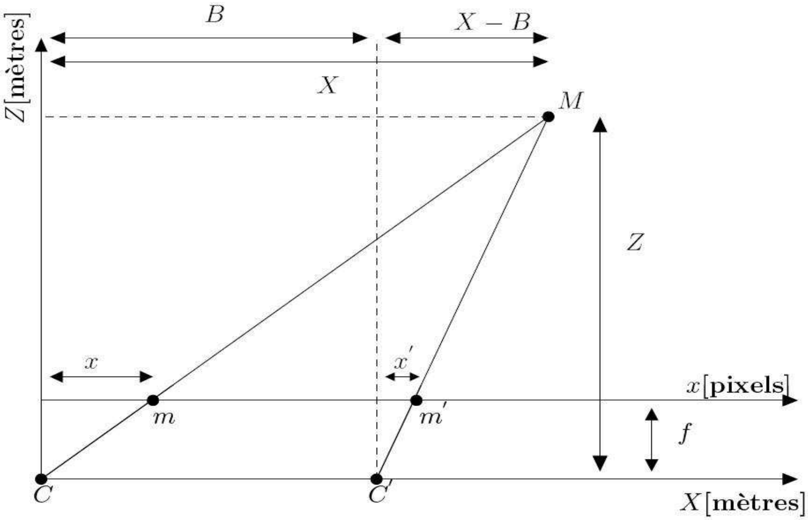
Left image (I_L)



Right image (I_R)

$$I_L(x, y) = I_R(x - u(x, y), y)$$

3D reconstruction



3D reconstruction

Objective

$$u(x,y) = x - x' = \frac{Bf}{Z}$$

Applications

- 3D television, 3D teleconferencing,
- Obstacle detection,
- Robotics, satellite, ...

Problem formulation

Objective

Find for each pixel in the left image I_L a corresponding pixel in the right image I_R.

State of the art

- Feature-matching [Medioni, Nevatia, 1985],
- Global method (dynamic programming [Veksler, 2002], variational approach [Deriche, Kornprobst,Aubert,1995,]),
- Normalized cross correlation [Zabih, Woodfill, 1994] . . .

Problem formulation

Objective

Find for each pixel in the left image I_L a corresponding pixel in the right image I_R.

State of the art

- Feature-matching [Medioni, Nevatia, 1985],
- Global method (dynamic programming [Veksler, 2002], variational approach [Deriche, Kornprobst,Aubert,1995,]),
- Normalized cross correlation [Zabih, Woodfill, 1994] ...

Variational method

$$J(u) = \sum_{(x,y)\in D} \phi(I_L(x,y) - I_R(x - u(x,y), y))$$

ϕ is assumed to belong to $\Gamma_0(\mathbb{R})$ which is the class of a proper lower-semi continuous convex function.

Convex minimization

- J is nonconvex with respect to the displacement field u .
- 1st order Taylor expansion of the nonlinear term around an initial estimate $\bar{u}$:

$$\boxed{I_R(x-u(x,y),y) = I_R(x-\bar{u}(x,y),y) - (u(x,y)-\bar{u}(x,y))I_R^x(x-\bar{u}(x,y),y)}$$

  * where I_R^x is the horizontal gradient of the disparity compensated right image.

Cost function

$$J(u) = \sum_{(x,y)\in\mathcal{D}} \phi(T(x,y)\ u(x,y) - r(x,y))$$

* $T(x,y) = I_R^x(x-\bar{u}(x,y),y)$
* $r(x,y) = I_R(x-\bar{u}(x,y),y) + \bar{u}(x,y)\ T(x,y) - I_L(x,y)$

Set theoretic estimation

- The minimization of functional J is an ill-posed problem.
- Additional **constraints** are required to regularize the solution.
- Formulate available constraints as closed convex sets in a Hilbert space $\mathcal{H}$:

Set theoretic estimation

- The minimization of functional J is an ill-posed problem.
- Additional **constraints** are required to regularize the solution.
- Formulate available constraints as closed convex sets in a Hilbert space $\mathcal{H}$:
- Admissibility problem

  Obtain a **feasible** solution minimizing an **objective** function and satisfying all **constraints** arising from prior knowledge.

Formulation

$$\text{Find } u \in S = \bigcap_{i=1}^{m} S_i \text{ such that } J(u) = \inf J(S).$$

Convex Constraints

Range Constraint

$$S_1 = \{u \in \mathcal{H} \mid u_{\min} \leq u \leq u_{\max}\}.$$

Convex Constraints

Range Constraint

$$S_1 = \{u \in \mathcal{H} \mid u_{\min} \leq u \leq u_{\max}\}.$$

Total variation Constraint

$$S_2 = \{u \in \mathcal{H} \mid \mathrm{TV}(u) \leq \tau\}.$$

$\tau > 0$

Convex Constraints

Range Constraint

$$S_1 = \{u \in \mathcal{H} \mid u_{\min} \leq u \leq u_{\max}\}.$$

Total variation Constraint

$$S_2 = \{u \in \mathcal{H} \mid \mathrm{TV}(u) \leq \tau\}.$$

$\tau > 0$

Wavelet Constraint

$$S_3 = \left\{ u \in \mathcal{H} \mid \sum_{j \geq 1, k \in \mathbb{Z}^2, o \in \{H,V\}} |c_{j,k,o}^{\mathcal{B}}| \leq \kappa \right\}$$

$\kappa > 0$, o is the orientation parameter and $j \in \mathbb{N}$ the resolution level.

Previous work

Subgradient Projections

- Discard occlusion areas $\mathcal{O}$.
- Quadratic criterion, Strictly convex: [Miled, Pesquet, Parent, 2009]

$$J(u) = \sum_{(x,y)\in\mathcal{D}\setminus\mathcal{O}} [T(x,y)u(x,y) - r(x,y)]^2 + \alpha \sum_{(x,y)\in\mathcal{D}} [u(x,y) - \bar{u}(x,y)]^2$$

Originality

Relax the strict convexity and the quadratic form of the function ϕ.

Proximity operator

Projection

The projection $P_C y$ of a point $y \in \mathbb{R}^N$ onto C is the solution to the problem:

$$\underset{x \in \mathbb{R}^N}{\text{minimize}} \quad \iota_C(x) + \frac{1}{2}\|x - y\|^2$$

where $\iota_C \in \Gamma_0(\mathbb{R}^N)$ is the indicator function of C.

Proximity operator

Projection

The projection $P_C y$ of a point $y \in \mathbb{R}^N$ onto C is the solution to the problem:

$$\underset{x\in\mathbb{R}^N}{\text{minimize}} \quad \iota_C(x) + \frac{1}{2}\|x-y\|^2$$

where $\iota_C \in \Gamma_0(\mathbb{R}^N)$ is the indicator function of C.

$\text{prox}_f\, y$

We replace the function ι_C by an arbitrary function $f \in \Gamma_0(\mathbb{R}^N)$. Then, the problem can be rewritten as

$$\underset{x\in\mathbb{R}^N}{\text{minimize}} \qquad f(x) + \frac{1}{2}\|x-y\|^2$$

This problem admits a unique solution which is the proximity operator $\text{prox}_f\, y$ of f at y.

Proposed approach

Optimization problem

$$\min_{L_i u \in C_i, i \in \{1,\dots,m\}} J(u) = \min_{L_i u \in C_i, i \in \{1,\dots,m\}} \sum_{(x,y) \in \mathcal{D} \setminus \mathcal{O}} \phi(T(x,y)\, u(x,y) - r(x,y))$$

Proposed approach

Optimization problem

$$\min_{L_i u \in C_i, i \in \{1,\ldots,m\}} J(u) = \min_{L_i u \in C_i, i \in \{1,\ldots,m\}} \sum_{(x,y)\in\mathcal{D}\setminus\mathcal{O}} \phi(T(x,y)\, u(x,y) - r(x,y))$$

- Each (S_i) can be expressed as $L_i^{-1}(C_i)$ where C_i is a non-empty closed convex subset of $\mathbb{R}^{N_i}$ and L_i is a matrix in $\mathbb{R}^{N_i \times K}$.
- Parallel proximal algorithm (e.g. PPXA+) allows us to minimize a convex criterion J on some closed convex constraint sets $(C_i)_{1\leq i\leq m}$.
- It consists of computing, in parallel, the projections onto the different convex sets $(C_i)_{1\leq i\leq m}$ and the proximity operator of the criterion J.

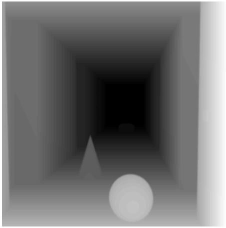
PPXA+ algorithm

Initialization $(\omega_1, ..., \omega_m) \in]0, +\infty[^m, \gamma > 0$
$(z_{i,0})_{1 \leq i \leq m+1} \in \mathbb{R}^{N_1} \times \ldots \times \mathbb{R}^{N_m} \times \mathbb{R}^K$
$Q = (\sum_{i=1}^m \omega_i L_i^\top L_i + \gamma \mathbb{I})^{-1}$
$u_0 = Q(\sum_{i=1}^m \omega_i L_i^\top z_{i,0} + \gamma z_{m+1,0})$
For $n = 0, 1, \ldots$ **do**
- **For** $i = 1, \ldots, m$ **do**
  - $p_{i,n} = P_{C_i}(z_{i,n})$
- **end For**
- $p_{m+1,n} = \text{prox}_{\frac{J}{\gamma}}(z_{m+1,n})$
- $c_n = Q(\sum_{i=1}^m \omega_i L_i^\top p_{i,n} + \gamma p_{m+1,n})$
- **For** $i = 1, \ldots, m$ **do**
  - $z_{i,n+1} = z_{i,n} + \lambda_n(L_i(2c_n - u_n) - p_{i,n})$
- **end For**
- $z_{m+1,n+1} = z_{m+1,n} + \lambda_n(2c_n - u_n - p_{m+1,n})$
- $u_{n+1} = u_n + \lambda_n(c_n - u_n)$

end For

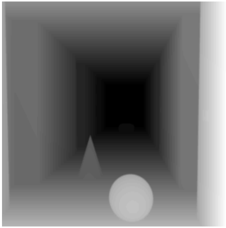
λ_n is a relaxation parameter.

Left image

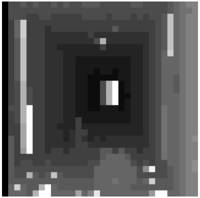


Ground truth

Left image

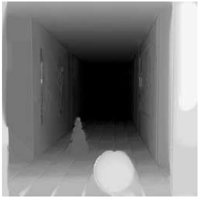


Ground truth



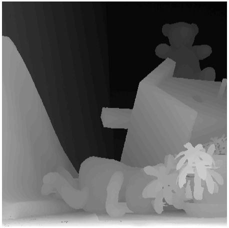
PSNR= 34.20 dB
Block BDE

PSNR= 34.83 dB
ℓ^2-norm DDE
subgradient projection



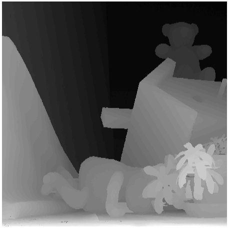
PSNR= 35.23 dB
ℓ^1-norm DDE
PPXA+ algo

Left image

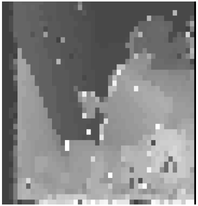


Ground truth

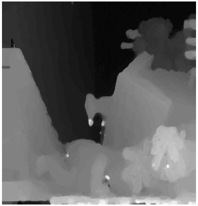
Left image



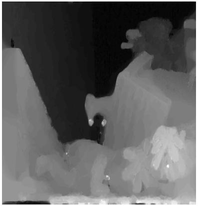
Ground truth



PSNR= 30.10 dB
Block BDE



PSNR= 37.08 dB
ℓ^2-norm DDE
subgradient projection



PSNR= 37.39 dB
ℓ^1-norm DDE
PPXA+ algo

Performances of the proposed method in stereo image coding

- Independent scheme: encoding separately the original images I_L and I_R by applying a 5/3 wavelet-like transform.

Performances of the proposed method in stereo image coding

- Independent scheme: encoding separately the original images I_L and I_R by applying a 5/3 wavelet-like transform.
- Joint coding scheme: applying the same transform to I_R and I_e, where:

$$\boxed{I_e(x, y) = I_L(x, y) - I_R(x - u, y)}$$

  * the resulting wavelet coefficients are encoded using JPEG2000 entropy codec.

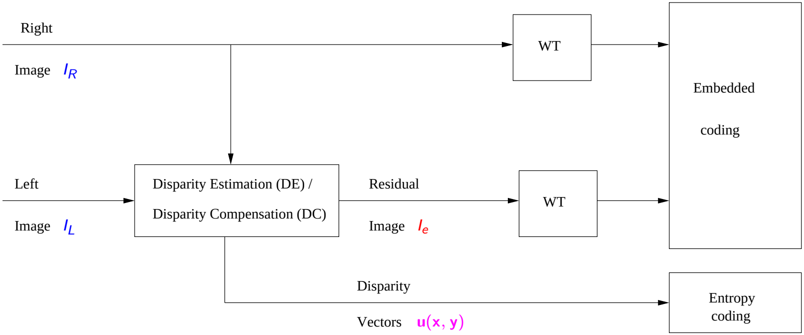
Performances of the proposed method in stereo image coding

- Independent scheme: encoding separately the original images I_L and I_R by applying a 5/3 wavelet-like transform.
- Joint coding scheme: applying the same transform to I_R and I_e, where:

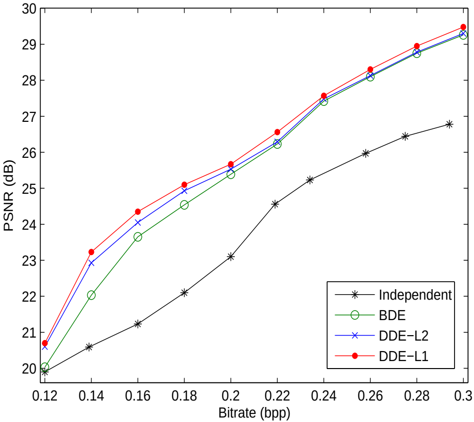
$$\boxed{I_e(x, y) = I_L(x, y) - I_R(x - u, y)}$$

  * the resulting wavelet coefficients are encoded using JPEG2000 entropy codec.
- The generated dense fields are encoded by applying a quadtree decomposition followed by an entropy coding with **H264/AVC** software.

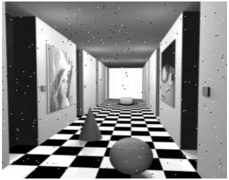
Performances of the proposed method in stereo image coding



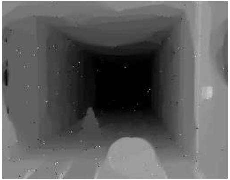
Performances of the proposed method in stereo image coding



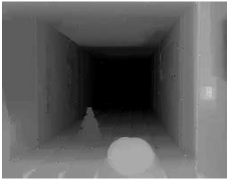
Noisy stereo pairs



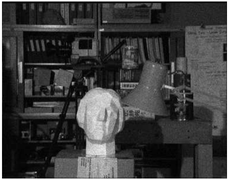
Corridor-Salt and paper



$C_1 + C_3$: SNR=15.19 dB, MAE=0.58
$\ell_2 - norm$ (*subgradient projection*)



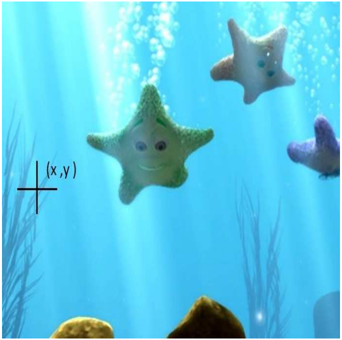
$C_1 + C_3$: SNR=15.68 dB, MAE= 0.5
PPXA+, $\ell_1 - norm$



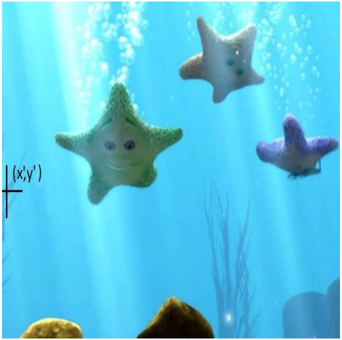
Tsukuba-Poisson

$C_1 + C_2$: SNR=17.76 dB, MAE=0.45
$\ell_2 - norm$ (*subgradient projection*)

$C_1 + C_2$: SNR=18.86 dB, MAE=0.46
PPXA+, *Kullback distance*



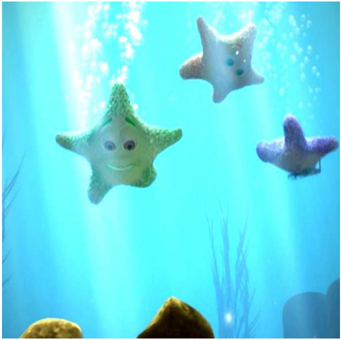
Left image (I_L)



Right image (I_R)

$$I_L(x, y) = I_R(x - u(x, y), y)$$

Left image (I_L)



Right image (I_R)

$$v(x, y) I_L(x, y) = I_R(x - u(x, y), y)$$

Cost function

$$J(u,v) = \sum_{(x,y)\in\mathcal{D}} \phi(T_1(x,y)\ u(x,y) + T_2(x,y)\ v(x,y) - r(x,y))$$

* $T_1(x,y) = I_R^x(x-\bar{u}(x,y),y),\ \ T_2(x,y) = I_L(x,y),$
* $r(x,y) = I_R(x-\bar{u}(x,y),y) + \bar{u}(x,y)\ T_1(x,y)$

Cost function

$$J(u,v) = \sum_{(x,y)\in\mathcal{D}} \phi(T_1(x,y)\ u(x,y) + T_2(x,y)\ v(x,y) - r(x,y))$$

- $T_1(x,y) = I_R^x(x - \bar{u}(x,y), y),\ \ T_2(x,y) = I_L(x,y),$
- $r(x,y) = I_R(x - \bar{u}(x,y), y) + \bar{u}(x,y)\ T_1(x,y)$

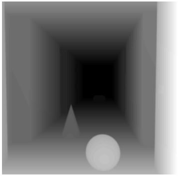
$w(x,y) = [u(x,y)\ v(x,y)]^\top$, $T(x,y) = [T_1(x,y)\ T_2(x,y)]$,

$$J(w) = \sum_{(x,y)\in\mathcal{D}} \phi(T(x,y)\ w(x,y) - r(x,y))$$

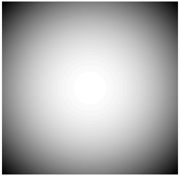
results

Left image

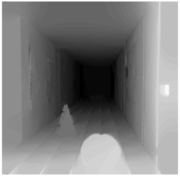
Right image



Disparity Ground truth



Illumination Ground truth

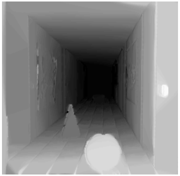


PSNR= 48.79 dB

Subgradient projection

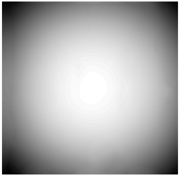
PSNR= 79.06 dB

Subgradient projection



PSNR= 55.83 dB

PPXA+ algo



PSNR= 87.26 dB

PPXA+ algo

Conclusion and perspectives

- Proposition of an efficient proximal method dealing with dense disparity estimation problems.
  - Direct projections[1] and proximity operators
  - Various criteria
  - Robustness w.r.t. perturbations
  - Color images
  - Images under illumination variation

[1] http://www.cs.ubc.ca/labs/scl/spgl1/download.html

Conclusion and perspectives

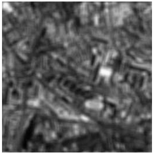
- Proposition of an efficient proximal method dealing with dense disparity estimation problems.
  - Direct projections[1] and proximity operators
  - Various criteria
  - Robustness w.r.t. perturbations
  - Color images
  - Images under illumination variation
- Good results w.r.t. existing works

Perspectives:

- Incorporate additional convex constraints.
- Parallel implementation (GPU).

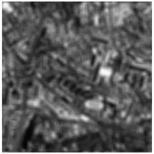
[1] http://www.cs.ubc.ca/labs/scl/spgl1/download.html

Context: solving inverse problems

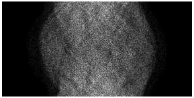


$\xrightarrow{?}$

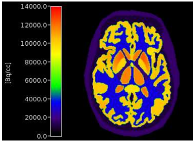
Context: solving inverse problems



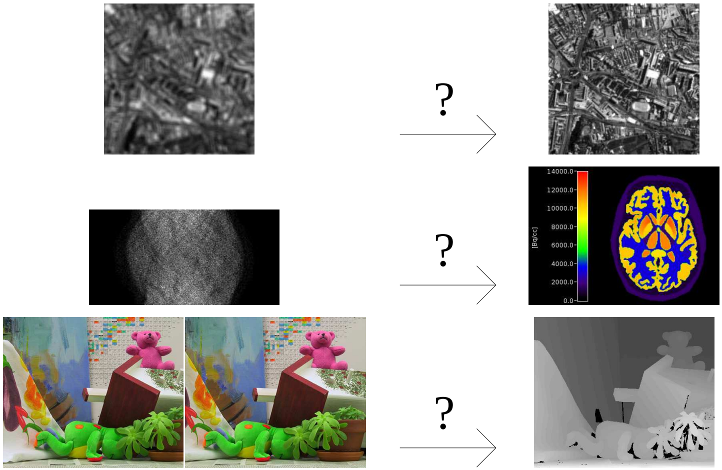
$\xrightarrow{?}$



$\xrightarrow{?}$



Context: solving inverse problems



Frame and convex optimization

- Quadratic regularization techniques (Wiener filtering).
- Multiresolution analyses used for denoising ($H = \mathrm{Id}$).
- Redundant frame repres. used for denoising.
- Forward-backward when $H \neq \mathrm{Id}$ [Combettes&Wajs 2005, Daubechies et al. 2004, Figueiredo&Bioucas-Dias 2003, Bect et al. 2004] $\rightarrow$ thresholded Landweber to solve $\|H \cdot -z\|_2^2 + \|\cdot\|_1$.

Frame and convex optimization

- Quadratic regularization techniques (Wiener filtering).
- Multiresolution analyses used for denoising ($H = \mathrm{Id}$).
- Redundant frame repres. used for denoising.
- Forward-backward when $H \neq \mathrm{Id}$ [Combettes&Wajs 2005, Daubechies et al. 2004, Figueiredo&Bioucas-Dias 2003, Bect et al. 2004] $\rightarrow$ thresholded Landweber to solve $\|H \cdot -z\|_2^2 + \| \cdot \|_1$.
- Douglas-Rachford (DR) algorithm [Combettes&Pesquet 2007]
- PPXA [Combettes&Pesquet 2008, Pustelnik et al. 2011]
- ADMM (SDMM) [Afonso et al., Setzer et al., Attouch & Soueycatt, 2009]
- Primal-Dual Algo. [Chen&Teboulle 1994, Esser et al. 2010, Combettes et al. 2011, Chambolle & Pock 2011, Briceño-Arias&Combettes 2011]
- PPXA+: unifying framework for PPXA and ADMM [Pesquet & Pustelnik 2011]

Publications

- Conf. • M. El Gheche, C. Chaux, J.-C. Pesquet, J. Farah and B. Pesquet-Popescu, Disparity map estimation under convex constraints using proximal algorithms, in SIPS 2011 , Beirut, Lebanon, 4-7 Oct. 2011.
- Conf. • M. El gheche, J.-C. Pesquet, C. Chaux, J. Farah et B. Pesquet-Popescu, Méthodes proximales pour l'estimation du champ de disparité à partir d'une paire d'images stéréoscopiques en présence de variations d'illumination, GRETSI 2011, Bordeaux, France, 5-8 sept. 2011.
- Conf. • M. El Gheche, J.-C. Pesquet, J. Farah, M. Kaaniche and B. Pesquet-Popescu, Proximal splitting methods for depth estimation, in ICASSP, Prague, Czech republic, 22-27 May 2011.
- Journal • C. Chaux, M. El Gheche, J. Farah, J.-C. Pesquet, and B. Pesquet-Popescu, A parallel proximal splitting method for disparity estimation from multicomponent images under illumination variation, accepted for publication in JMIV, Springer Netherlands, 2012.

Thank you !