

Proximal splitting methods for depth estimation.

Mireille EL GHECHE

Joint work with J.-C. Pesquet¹, J. Farah², Mounir Kaaniche^{1,3}
and Béatrice Pesquet-Popescu³

¹Laboratoire d'Informatique Gaspard Monge - UMR CNRS 8049, Université Paris-Est, France.

²Departement of Telecommunications, Faculty of Engineering, Holy-Spirit University of Kaslik, Lebanon.

³Département Traitement du Signal et des Images, Telecom-ParisTech, France.

ICASSP 2011, Prague, 23-27 May 2011



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)

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

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} :

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

- * where ∇I_d^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

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

$$\tau > 0$$

Convex Constraints

Range Constraint

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

Total variation Constraint

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

$\tau > 0$

Wavelet Constraint

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

$\kappa > 0$, 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

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

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

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

- ▶ 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 $(\omega_1, \dots, \omega_m) \in]0, +\infty[^m, \gamma > 0$

$(z_{i,0})_{1 \leq i \leq m+1} \in \mathbb{R}^{N_1} \times \dots \times \mathbb{R}^{N_m} \times \mathbb{R}^K$

$Q = (\sum_{i=1}^m \omega_i L_i^\top L_i + \gamma \mathbb{I})^{-1}$

$u_0 = Q(\sum_{i=1}^m \omega_i L_i^\top z_{i,0} + \gamma z_{m+1,0})$

For $n = 0, 1, \dots$ **do**

For $i = 1, \dots, m$ **do**

$p_{i,n} = P_{C_i}(z_{i,n})$

end For

$p_{m+1,n} = \text{prox}_{\frac{\gamma}{\omega_{m+1}}}(z_{m+1,n})$

$c_n = Q(\sum_{i=1}^m \omega_i L_i^\top p_{i,n} + \gamma p_{m+1,n})$

For $i = 1, \dots, m$ **do**

$z_{i,n+1} = z_{i,n} + \lambda_n(L_i(2c_n - u_n) - p_{i,n})$

end For

$z_{m+1,n+1} = z_{m+1,n} + \lambda_n(2c_n - u_n - p_{m+1,n})$

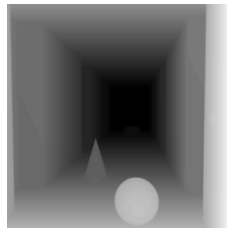
$u_{n+1} = u_n + \lambda_n(c_n - u_n)$

end For

λ_n is a relaxation parameter.



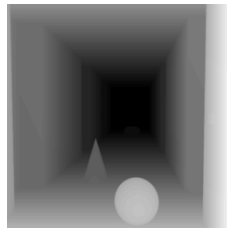
Left image



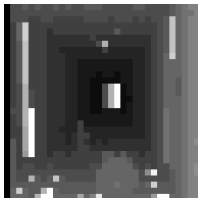
Ground truth



Left image



Ground truth



PSNR= 34.20 dB

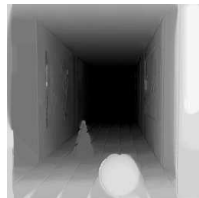
Block BDE



PSNR= 34.83 dB

ℓ^2 -norm DDE

subgradient projection



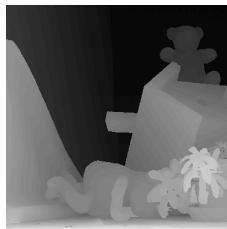
PSNR= 35.23 dB

ℓ^1 -norm DDE

PPXA+ algo



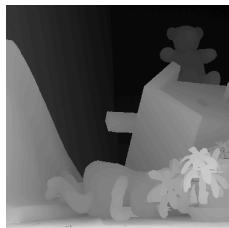
Left image



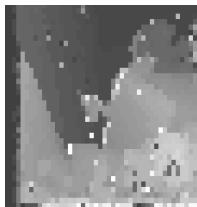
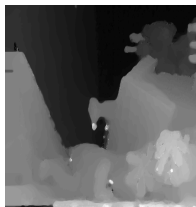
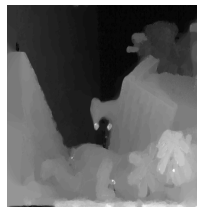
Ground truth



Left image



Ground truth

PSNR= 30.10 dB
Block BDEPSNR= 37.08 dB
 ℓ^2 -norm DDE
subgradient projectionPSNR= 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:

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

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

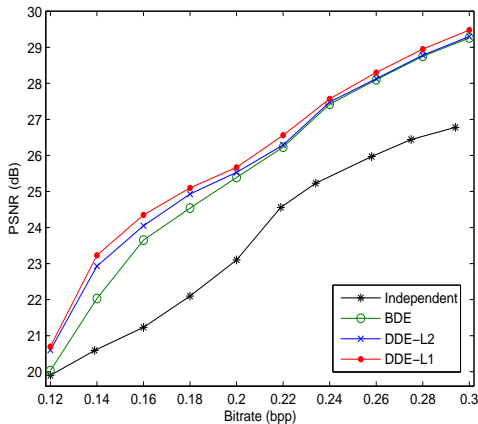
Performances of the proposed method in stereo image coding

- ▶ **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:

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

- * 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



Noisy stereo pairs



Corridor-Salt and paper

 $C_1 + C_3$: SNR=15.19 dB, MAE=0.58
 ℓ_2 - norm (subgradient projection) $C_1 + C_3$: SNR=15.68 dB, MAE= 0.5
PPXA+, ℓ_1 - norm

Tsukuba-Poisson

 $C_1 + C_2$: SNR=17.76 dB, MAE=0.45
 ℓ_2 - norm (subgradient projection) $C_1 + C_2$: SNR=18.86 dB, MAE=0.46
PPXA+, Kullback distance

Conclusion and perspectives

- ▶ Proposition of an efficient proximal method dealing with dense disparity estimation problems.
 - ▶ Direct projections¹ and proximity operators
 - ▶ Various criteria
 - ▶ Application in presence of perturbations.

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 projections¹ and proximity operators
 - ▶ Various criteria
 - ▶ Application in presence of perturbations.
- ▶ Enhanced results w.r.t. existing works

Perspectives:

- ▶ Extension of the approach to the case where illumination variations exist.
- ▶ Take into account color images.

¹<http://www.cs.ubc.ca/labs/scl/spg1/download.html>

Thank you !