Disparity map estimation under convex constraints using proximal algorithms
Mireille El Gheche, Jean-Christophe Pesquet, Youmana Farah, Caroline Chaux, Béatrice Pesquet-Popescu

To cite this version:

HAL Id: hal-00733655
https://hal-upec-upem.archives-ouvertes.fr/hal-00733655
Submitted on 19 Sep 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
DISPARITY MAP ESTIMATION UNDER CONVEX CONSTRAINTS USING PROXIMAL ALGORITHMS

Mireille El Gheche¹,², Jean-Christophe Pesquet¹, Joumana Farah², Caroline Chaux¹ and Béatrice Pesquet-Popescu³

¹ Université Paris-Est LIGM and UMR-CNRS 8049, 77454 Marne-la-Vallée cedex, France
{el.gheche,pesquet,chaux}@univ-mlv.fr

² Department of Telecommunications, Faculty of Engineering, Holy-Spirit University of Kaslik, P.O. Box 446, Joumieh, Lebanon
joumanafarah@usek.edu.lb

³ Telecom ParisTech, Signal and Image Proc. Dept. 75014 Paris, France
pesquet@telecom-paristech.fr

ABSTRACT

In this paper, we propose a new approach for estimating depth maps of stereo images which are prone to various types of noise. This method, based on a parallel proximal algorithm, gives a great flexibility in the choice of the constrained criterion to be minimized, thus allowing us to take into account different types of noise distributions. Our main objective is to present an iterative estimation method based on recent convex optimization algorithms and proximal tools. Results for several error measures demonstrate the effectiveness and robustness of the proposed method for disparity map estimation even in the presence of perturbations.

Index Terms— Proximity operator, total variation, tight frame, proximal algorithms, convex optimization, ℓ_p-norm, Kullback-Leibler divergence.

1. INTRODUCTION

Given any two images of the same scene acquired by stereoscopic cameras, it can be seen that a degree of similarity exists between the two views. Indeed, the pixel at position x in the left image I_L corresponds to a pixel at position x – u in the right image I_R: the disparity of those pixels is denoted by u. The matching problem amounts to searching for the disparity field u which minimizes an error measure. Consequently, we can express the stereo matching problem as:

\[ \text{minimize } J(u), \]  

where,

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

and \( \phi \) is assumed to belong to \( \Gamma_0(\mathbb{R}) \) which is the class of a proper lower-semi continuous convex function from \( \mathbb{R} \) to \( [-\infty, +\infty] \). \( D \subset \mathbb{Z}^2 \) is the considered finite image domain.

The recovery of the depth estimation has been increasingly used in a variety of fields such as movies, web networks, 3D reconstruction and computer games. During the last decades, various methods have been introduced in disparity estimation, especially in computer vision [1]. Stereo matching algorithms are generally classified into two categories: feature based and area based ones. Feature based methods match the feature elements between two images [2], while area based methods perform matching between pixels. Currently, the associated problems are often solved using discrete optimization techniques based on a global approach. Algorithms such as dynamic programming [3], graph cuts [4] and variational methods [5] were proposed. Note that variational-based disparity estimation methods demonstrated excellent performance compared with the state-of-the-art. In [5], a convex energy function approximation of J is derived and minimized subject to various convex constraints (modeled by convex sets \( (C_i)_{1 \leq i \leq m} \) arising from prior knowledge and observed data. The optimization problem (1) to be solved then becomes

\[ \text{minimize } \sum_{L_i \in C_i, i \in \{1, \ldots, m\}} J(u), \]  

where \( (L_i)_{1 \leq i \leq m} \) are linear operators. However, the employed algorithm requires not only the strict convexity of the criterion to be considered but a quadratic form of it.

This paper describes a new DDE (Dense Disparity Estimation) approach that generalizes the convex optimization approximation presented in [5]. The proposed method is based on some proximity operator (prox) properties and then, it is no longer limited to strictly convex quadratic criteria. Therefore, explicit expressions of the prox of some norms such as \( \ell_1 \)-norm and more generally \( \ell_p \)-norm with \( p > 1 \) or the Kullback-Leibler divergence allow us to apply such proximal algorithms to our disparity map estimation problem. By allowing a rich choice of distance functions, the proposed technique is well-suited for dealing with different types of noise corrupting the observed data, such as Poisson noise,
salt and pepper noise or Gaussian noise.

The remainder of the paper is organized as follows: Section 2 describes the basis tools employed in our approach. Section 3 lays out the problem statement. Results on the Middlebury dataset\footnote{http://cat.middlebury.edu/stereo/newdata.html} are provided in Section 4. Finally, some conclusions are drawn in Section 5.

2. BACKGROUND

We will now present some tools which are useful in the solution of our minimization problem. In what follows, we will focus on the definition of the proximity operators, total variation and tight frames.

2.1. Proximity operator

The proximity operator of a convex function is a natural extension of the projection operator $P_C$ onto a nonempty closed convex set $C \subset \mathbb{R}^N$ [6, 7, 8]. The projection $P_{Cy}$ of a point $y \in \mathbb{R}^N$ onto $C$ is the solution to the problem:

$$\minimize_{x \in \mathbb{R}^N} \ i_C(x) + \frac{1}{2} \|x - y\|^2 \quad (4)$$

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

In a seminal paper [9], Moreau proposed the following extension of the notion of projection by replacing the function $i_C$ by an arbitrary function $f \in \Gamma_0(\mathbb{R}^N)$. Then, the problem can be rewritten as

$$\minimize_{x \in \mathbb{R}^N} \ f(x) + \frac{1}{2} \|x - y\|^2 \quad (5)$$

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

Proximity operators have very attractive properties [7] that make them particularly well-suited to design iterative minimization algorithms. In Table 1, we summarize some of them that will be needed to obtain closed form expressions of the required proximity operators.

2.2. Total variation

Let $H$ and $V$ be two spatial convolution operators corresponding respectively to the computation of horizontal and vertical discrete gradients. Then, a discrete version of the total variation (TV) for which proximity tools can be applied is the following [10, 11]:

$$\text{TV}(u) = \sum_{(x,y) \in D} \sqrt{|H u(x,y)|^2 + |V u(x,y)|^2} \quad (6)$$

It should be noticed that, in our simulations, we apply periodic convolutions, which means that the operators $H$ and $V$ are diagonalized by a Fast Fourier Transform.

<table>
<thead>
<tr>
<th>$f(x)$</th>
<th>$\text{prox}_f x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi(x - z)$, $z \in \mathbb{R}^N$</td>
<td>$z + \text{prox}_\phi(x - z)$</td>
</tr>
<tr>
<td>$\phi(x/\rho)$, $\rho \in \mathbb{R} \setminus {0}$</td>
<td>$\rho \text{ prox}_\phi(x/\rho)$</td>
</tr>
<tr>
<td>$\phi(Lx)$, semi-orthogonal, $L \in \mathbb{R}^{M \times N}$, $LL^T = \nu I$, $\nu &gt; 0$</td>
<td>$x + \nu^{-1}L^T(\text{prox}_\varphi(Lx) - Lx)$</td>
</tr>
<tr>
<td>$\mu_C(x) = \begin{cases} 0 &amp; \text{if } x \in C \ +\infty &amp; \text{if } x \notin C \end{cases}$</td>
<td>$P_{C \mu} x$</td>
</tr>
<tr>
<td>$\chi &gt; 0$, $\alpha &gt; 0$</td>
<td>$\begin{cases} -\chi \ln(x) + \alpha x, &amp; \text{if } x &gt; 0 \ +\infty &amp; \text{if } x \leq 0 \end{cases}$</td>
</tr>
</tbody>
</table>

Table 1. Some proximity operator properties ($I$ denotes the identity matrix) [7].

2.3. Tight frame

Frame representations [12] and more precisely tight frame representations have been widely used during the last decade. Such transforms can be described by an analysis frame operator $F$ and a synthesis frame operator $F^T$. $F$ is said to be a tight frame when $F F^T = \nu I$, where $\nu > 0$.

A simple example of a tight frame is the union of $m$ orthonormal Haar bases over 1 resolution level.

In this paper, we will restrict our attention to the union of $m$ shifted orthonormal Haar bases over 1 resolution level.

2.4. PXXA+ Algorithm

Many algorithms have been formulated in the literature to solve optimization problems like (1). In our case, we use the Parallel Proximal Algorithm (PXXA+) [14] which is a flexible tool. PXXA+ (algorithm 1) 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 $J$.

Suppose that the following assumptions hold.

- $\sum_{i=1}^m L_i^T L_i + I$ is an invertible matrix.
- $\exists \ u \in \mathbb{R}^K$ such that $(\forall i \in \{1, \ldots, m\}) \ L_i u \in \rho i(C_i)$ and $u \in \rho i(\text{dom } J)$\footnote{\text{dom } J \text{ is the domain of } J \text{ and the relative interior of a set } C \text{ is denoted by } \rho i C}.
- $(\lambda_n)_{n \in \mathbb{N}}$ is a sequence of relaxation parameters such that $(\forall n \in \mathbb{N}) \ \lambda \leq \lambda_{n+1} \leq \lambda_n < 2$, where $\lambda \in [0, 2]$.

Then, the sequence $(u_n)_{n \in \mathbb{N}}$ generated by Algorithm 1 converges to a solution to Problem (3), provided that such a solution exists.
Based on (2) and (7), we deduce that:

\[ J(u) = \sum_{(x,y) \in D \setminus \mathcal{O}} \phi(T(x,y) u(x,y) - r(x,y)) \]  

The proposed PPXA+ algorithm provides an efficient solution to solve this problem. Our main contribution here lies in the various choices that can be made for the error measure. As already mentioned, the algorithm is based on iterating computations of the proximity operator of \( J \) and projections onto convex sets related to constraints. We will first give the explicit forms of the proximity operators of interest in this work which correspond to error measures \( J \) taken as \( \ell_p \)-norms or Kullback-Leibler divergence.

### 3.2. Error measures

#### 3.2.1. \( \ell_p \)-norm

Based on Table 1, it can be deduced that

\[ \tilde{u}(s) = \text{prox}_{\gamma^{-1}|T(s)|^2 \phi} (T(s)\tilde{u}(s) - r(s)) + r(s) \]

where \( s = (x,y) \). Then the proximity operator of \( \gamma^{-1}|T(s)|^2 \phi \) with \( \phi = | \cdot |^p \) and \( \xi(s) = T(s)\tilde{u}(s) - r(s) \), is [15]:

\[
\text{prox}_{\gamma^{-1}|T(s)|^2 \phi} (\xi(s)) =
\begin{cases}
\text{sign}(\xi(s)) \max(\xi(s) - \gamma^{-1}|T(s)|^2, 0), & \text{if } p = 1; \\
\frac{\xi(s)}{2\gamma^{-1}|T(s)|^2 + 1}, & \text{if } p = 2; \\
\text{sign}(\xi(s)) \sqrt{1 + 12\gamma^{-1}|T(s)|^2|\xi(s)| - 1}, & \text{if } p = 3; \\
\left( \frac{\mu + \xi(s)}{8\gamma^{-1}|T(s)|^2} \right)^\frac{1}{2} - \left( \frac{\mu - \xi(s)}{8\gamma^{-1}|T(s)|^2} \right)^\frac{1}{2}, & \text{if } p = 4,
\end{cases}
\]

where \( \mu = \sqrt{\xi(s)^2 + \frac{1}{2\gamma^{-1}|T(s)|^2}} \).  

#### 3.2.2. Kullback-Leibler divergence

One can note that (7) can be rewritten as

\[ I_R(x - u(s), s) = r(s) + I_L(s) \cdot u(s) \cdot T(s) = \tilde{r}(s) - u(s) T(s) = \xi(s). \]

Then, the criterion to be minimized takes a form slightly different from (11):

\[ J(u) = \sum_{(x,y) \in D \setminus \mathcal{O}} \Phi(I_L(s), \xi(s)) \]

where \( \Phi(I_L(s), \xi(s)) \) is the Kullback-Leibler divergence function given by:

\[ \Phi(I_L(s), \xi(s)) = \frac{1}{2} \sum_{(x,y) \in D \setminus \mathcal{O}} \left( I_L(s) \cdot u(s) \cdot T(s) - \xi(s) \right)^2. \]
The properties in Table 1 allow us to re-express the proxim-
ity operator \( \tilde{u}(s) \) at \( u(s) \) of the function \( \gamma^{-1}\Phi (I_L(s), \tilde{r}(s) - T(s) \cdot ) \) as:

\[
(\forall s \in D \setminus \mathcal{O})
\tilde{u}(s) = \frac{1}{T(s)} \left( \tilde{r}(s) - \text{prox}_{\gamma^{-1}|T(s)|^2\Phi (I_L(s), \cdot )} \left( \zeta (s) \right) \right)
\]

where \( \zeta (s) = \tilde{r}(s) - \tilde{u}(s) T(s) \). If \( I_L(s) \in [0, \infty ] \), then

\[
\text{prox}_{\gamma^{-1}|T(s)|^2\Phi (I_L(s), \cdot )} \left( \zeta (s) \right) = \frac{1}{2} \sqrt{\zeta(s)^2 - \gamma^{-1}|T(s)|^2}^2 + 4\gamma^{-1}|T(s)|^2 I_L(s)
\]

Furthermore, it can be shown that the above expression re-
mains valid if \( I_L(s) = 0 \). Let us now turn our attention to the
convex constraints that can be introduced.

### 3.3. Convex constraints

The motivation behind introducing these constraints is to in-
corporate additional prior knowledge about the solution.

First, we will constrain the range values of the disparity \( u \)
by setting a first constraint:

\[
S_1 = \{ u \mid u_{\text{min}} \leq u \leq u_{\text{max}} \}.
\]

Then, we can impose an upper bound on the total variation
measure. Indeed, this constraint allows us to recover piece-
wise and homogeneous areas with sharp edges:

\[
S_2 = \{ u \mid TV(u) \leq \tau_u \}.
\]

Finally, one can introduce some prior information on the
wavelet coefficients of the disparity. Based on Section 2.3,
we can represent the associated constraint in terms of Besov
space \( B_{1,1}^{2} \) semi-norm as:

\[
S_3 = \{ u \mid \sum_{j \geq 1, k \in \mathbb{Z}^2, a \in \{H, V\}} |c_{j,k,a}^{H}(u)| \leq \kappa \}.
\]

One can link these constraints \( S_1 \leq i \leq 3 \) and some convex
sets \( \{ C_i \}_{1 \leq i \leq 3} \) by noticing that \( S_1 = C_1 \), \( S_2 = L_2^{-1}(C_2) \) where
\( L_2 = \begin{bmatrix} H \\ V \end{bmatrix} \) and \( S_3 = L_3^{-1}(C_3) \) with \( L_3 = F \). The
projection onto \( C_1 \) is straightforward. The projections onto the
convex sets \( C_2 \) and \( C_3 \) were implemented using the iterative
projection algorithm proposed in [16].

### 4. SIMULATION RESULTS

The objective of this section is to present numerical results
obtained with the proposed method for a variety of synthetic
and noise images. In [17], we provided some comparisons
with the state-of-the-art DDE method presented in [5]. We
brieﬂy report the results of the application of our approach to
the standard Middlebury stereo dataset. The objective of this
first experiment is to demonstrate the validity of the proposed
method without any perturbation. Fig. 1 shows the stereo
image pairs considered in this work, namely Corridor, Saw,
Teddy and Cones, along with their ground truth images. Our
method allows us to compute a smooth disparity map with ac-
curate depth discontinuities as also shown by the preliminary
results in [17]. We specify in each case the considered error
measure and constraints (the choice was made according to
the best performances). We evaluate our method using two
error measures between the computed ﬁeld \( u_{c} \) and the ground
truth ﬁeld \( u_{\text{ref}} \):

\[
\text{SNR} = 10 \log_{10} \left( \frac{\sum_{s \in D} |u_{\text{ref}}(s)|^2}{\sum_{s \in D} |u_{c}(s) - u_{\text{ref}}(s)|^2} \right),
\]

\[
\text{MAE} = \frac{1}{N} \sum_{s \in D} |u_{c}(s) - u_{\text{ref}}(s)|,
\]

where \( N \) is the number of pixels.

In the presence of noise, we evaluate our method using dif-
f erent kinds of perturbation. To introduce a significant noise
variation, we modiﬁed the left and the right images by using the
imnoise function in Matlab. Fig. 2 shows the computed
disparity maps from the noisy stereo pairs. As we can see,
our method is less affected than DDE by noise changes and
provides accurate depth maps.

We display in Fig. 2 results obtained with the proposed
error measures (\( \ell_1 \)-norm for salt and paper noise, Kullback-
leibler divergence for Poisson noise and quadratic norm for
Gaussian noise). Different constraints were also taken into
account: we chose the one that leads to the best results ei-
ther for DDE or PPXA+. As shown in the last column, our
approach outperforms the DDE method [5], which proves the
effectiveness of our method even in the presence of Gaussi-
noise. Furthermore, as we can see in the first and second
columns, numerical values and visual comparisons on the ob-
tained disparity maps conﬁrm that our PPXA+ algorithm pro-
duces good estimates in the presence of other kinds of noise.

Finally, Fig. 3 displays the MAE with respect to the Pois-
son noise intensity \( \alpha \) (a smaller \( \alpha \) corresponding to a higher
noise). PPXA+ algorithm outperforms the DDE approach es-
pecially when the noise intensity is high.

### 5. CONCLUSIONS

An efficient proximal method that deals with disparity esti-
mation problems has been proposed in this paper. Being very
flexible, it allows us to minimize various criteria (not neces-
sarily differentiable nor strictly convex), which can be useful
when perturbations are present during image acquisition. The effectiveness of the proposed approach was demonstrated on different stereo image pairs, while taking into account different kinds of perturbations. In the future, we plan to extend our approach to the case where illumination variations exist.

6. REFERENCES


**Fig. 2.** From left to right: Salt and paper noise, Poisson noise, Gaussian noise. From top to bottom: noisy left image, DDE results, PPXA+ results.

![Fig. 2](image1)

**Fig. 3.** MAE with respect to Poisson noise intensity $\alpha$ applied to Cones pair.

![Fig. 3](image2)