Planar Patch Detection for Disparity Maps

Eric Bughin
CMLA - ENS Cachan
61, avenue du Président Wilson
94235 Cachan Cedex – FRANCE
bughin@cmla.ens-cachan.fr

Andrés Almansa
LTCI Telecom ParisTech
46, rue Barrault
75634 PARIS Cedex 13 – FRANCE
andres.almansa@telecom-paristech.fr

Abstract

We propose a new parameter-free method for detecting planar patches in disparity maps. We first introduce an a contrario *decision criterion which may be used to solve two decision problems on configurations of 3D points:* (i) *is the configuration well explained by a plane?;* (ii) *what is the optimal number of planes that best explains the configuration? These decision criteria are the core of an algorithm that searches for an optimal explanation of a disparity map by planar patches whenever applicable. This method may be used for 3D reconstruction of urban environments, particularly in the context of low-baseline stereo where precision requirements are most strict, and a pertinent choice of the type and amount of regularization is key to achieving accurate results. It also suggests its use for automatic vectorization of urban DEMs, where a sensible geometric representation is key to achieving good visualizations.*

1. Introduction

Finding a Vectorized Digital Elevation Model (VDEM) of a urban scene is of great interest. This is necessary for various applications such as 3D data compression, urban planning, radiowave reachability, disaster recovery, etc. This paper concentrates on a crucial step in the obtention of VDEMs, namely the optimal grouping of 3D point clouds representing the underlying surface, into planar patches, whenever possible.

1.1. Small baseline stereo

Even though the input 3D point clouds may come from a variety of sources (including LIDAR, SAR interferometry, etc.), and despite the potential applicability of our technique in those settings, we concentrate here on the specific case of disparity maps obtained by photogrammetry from low-baseline stereo pairs [4] (however, the method we propose here does still work in the case of large-baseline stereovision). Such 3D measurement systems have a certain number of advantages: *(i)* sure and independent punctual matches become feasible in a relatively dense area [15]; *(ii)* occlusions are reduced to a minimum and quasi-zenital views can be assumed. They however introduce new challenges, since fattening artifacts become specially important, and highly subpixel-accurate disparities are required to obtain a usable accuracy in height. For this reason careful regularization techniques (like the robust affine regression we propose here) are crucial to obtaining the required accuracy level. Furthermore, such a regularization should be appplied only after verifying that the underlying data is well explained by the chosen regularization model (thus enabling the use of other regularization or interpolation techniques for non planar structures such as vegetation, domes, etc), and level (thus avoiding over- of sub-regularization).

1.2. Previous work

Various methods were previously proposed to find a 3D model from a disparity map, but for different reasons they do not fit the low B/H setting.

RANSAC and Hough transform. Robust estimators based on RANSAC or Hough transform have been widely used in unconstrained stereovision.

Though the Hough transform seems pretty natural for the detection of multiple objects, its computational complexity makes it hard to use when the objects are defined by too many parameters (typically more than 2).

One of the challenges to overcome when using RANSAC is that it was originally designed to detect only one object among outliers. Though various approaches were proposed to overcome this problem, most of them [22, 23, 20] were proved to fail in the experiments of [19]. This is mainly due to the detection of phantom objects made of the combination of two objects. In [19], groups are merged according to their Jacquard distance. The main problem of this algorithm is to get rid of groups made only of outliers. This points out

the need of a criterion to decide whether a group is valid or not.

Such a decision criterion was addressed in [14] in the context of group matching of SIFT descriptors to find different objects in a scene. However, this technique does not scale well to quasi-dense correspondence maps, where transformations to be detected are far more numerous. In [12] k-nearest neighbors and local orientation were combined with RANSAC to fit models to dense 3D data, but several thresholds have to be fine-tuned.

Geometric modelling from disparity maps. Due to the larger number of both points and objects in disparity maps, other methods than RANSAC or Hough transform have been proposed.

In [13] the authors used a dictionary of complex building models to fit the disparity map. However the applicability to the low-baseline case is less evident because the initial delimitation of buildings by rectangle-fitting to the disparity map is more error prone when the latter is noisy and affected by fattening (adhesion) artifacts. In addition, the slow convergence of the underlying non-convex optimization procedure, may scale up when the number of models in the dictionary is increased to more closely fit reality. Another version of this algorithm drops the rectangle-fitting part but has the main drawback of requiring considerable user interaction.

In [2], the authors tried to match line segments of both images in order to find the height of a 3 dimensional edge. Then half planes are computed on each side of the segment. Despite their good results in urban areas, their method does not apply to low baseline stereo, because it relies on segment-to-segment matching, which proved to be not precise enough in this case [15]. Furthermore, when segments are badly or not detected, no assumption can be made.

Various methods were proposed for the segmentation of range images [11, 18, 9] but they all lack at some point of a generic criterion to decide when a group can be considered as planar.

In [10], the authors propose an *a contrario* region merging procedure to obtain a piecewise affine disparity map. However, the procedure is highly dependent on an initial partition which can be error prone. This initialisation is obtained by assuming that quasi-uniform gray-levels imply a common affine model, which is often, but not always the case, even under Lambertian hypotheses.

Following the same hypothesis, [7] uses the luminance-geodesic Voronoi cells of a sparse disparity map to provide an initial piecewise affine interpolation. From this partition, a merging procedure is proposed to find the final interpolation. However, due to the computational complexity of the geodesic distance, its use with our quasi-dense disparity maps becomes prohibitive.

1.3. Overview

The paper is organized as follows. In Section 2, we introduce an *a contrario* criterion [6] for validating a planar model and for selecting the best model among single or multiple planes. This criterion is similar to the one introduced in [10] but does not rely on the segmentation constraint. Unlike classical model selection criteria [1, 17] which are quite similar in nature, the proposed *a contrario* criterion serves also as an automatic and parameter free validation method. In Section 3, we propose an algorithm using the *a contrario* criterion to search for the different planes in a disparity map by means of a split & merge strategy. At last, the experimental results in Section 4 support the effectiveness of the proposed algorithm and its potential use for interpolation, denoising, and vectorization of urban DEMs.

2. *A contrario* plane validation

Most methods presented before lack of an automatic criterion to decide when a group found by any algorithm actually is a plane. When a decision is made, this is usually done by keeping the groups for which the size is superior to a predefined threshold, which is difficult to tune in a universal manner. We propose here to use an *a contrario* framework to make the decision automatic.

Following a similar methodology as the one done in [10], a group can be considered as planar if the probability of finding such a large number of points on a plane is very low if the disparity map was random (*background process*). In such a case, according to the Helmholtz principle, the background hypothesis has to be rejected in favor of the detection of a significant planar patch.

2.1. Data and background process

A (deterministic) disparity map z is a mapping where each point X in a discrete grid $\mathcal{D} \subset \mathbb{R}^2$ (image plane) is associated to a disparity (or height) $z \in \mathbb{R}$. z depends on the observed object on the disparity map and can take any value in $[z_{min}, z_{max}] \subset \mathbb{R}$. It can also be seen as a realization of the random process $\mathcal{Z}(X)$ defined as:

Definition 1 (Background process) *A background process is a finite process $\mathcal{Z}(X) \sim \mathcal{U}([z_{min}, z_{max}])$, $X \in \mathcal{D}$, made of mutually independent variables.*

Let N be the number of points in the discrete grid $\mathcal{D}$. For a given plane π and a group G of n data points, the *a contrario* criterion should consist of comparing the (random) number $\mathcal{K}$ of points near π out of N points drawn from the background process to the actual number k of points from G near π. If the probability $\mathbb{P}[\mathcal{K} \geqslant k]$ is small enough then the

planar grouping π of the points in G cannot be simply explained by the background process and a meaningful plane is detected.

However, such a comparison penalizes small groups (even when they are planar) if one set the value of N. Moreover, the final goal of such a criterion is to detect not only planes but planar facets which are very localized in the disparity map. Therefore a good comparison should be made on a region of the disparity map containing the tested data point group. For the rest of this article, let $\mathcal{R}$ be a sufficiently rich set of regions such that $\bigcup_{R\in\mathcal{R}} R = \mathcal{D}$.

2.2. Meaningful planes

Definition 2 (NFA or Number of False Alarms) *Let $G \subset \mathcal{D}$ be a group of points. Let $R \supset G$ be a region containing G, π be a plane and σ_z the z-tolerance to belong to π (σ_z may be different at each point). The NFA of G according to (R, π, σ_z) is defined as:*

$$NFA(G, R, \pi, \sigma_z) \equiv N_{tests}\mathbb{P}[\mathcal{K} \geqslant k] \tag{1}$$

where

- $k = \sum_{X\in G} \mathbb{1}_{\{|z(X)-z_\pi(X)|<\sigma_z(X)\}}$ counts the number of points from G that are sufficiently close to π.
- $\mathcal{K} = \sum_{X\in R} \mathbb{1}_{\{|\mathcal{Z}(X)-z_\pi(X)|<\sigma_z(X)\}}$ is a random variable counting the number of random points in R that are sufficiently close to plane π,
- N_{tests} is the number of tests done to go through all the planes in all the tested regions.

The probability $P(\mathcal{K} \geqslant k)$ is then the probability that a random vector $\mathcal{Z}(R)$ with size given by R shows at least as many points near π as what is observed by $z(G)$. The NFA is then a measure of how likely the closeness of $z(G)$ to π is to happen by chance.

Remark 1 *In most of the methods described during the introduction, the precision over the points σ_z is supposed to be known. For approaches such as RANSAC, the choice of planar groups directly depends on it and a bad value may lead to disastrous results. In [15], Sabater gave an estimation of the finest precision that can be expected at each point when a disparity map is computed using a block matching method. The parameter σ_z in the NFA computation can therefore be fixed to the value given by [15].*

However, in cases where the precision is unknown the NFA gives a proper framework to choose this precision parameter. The value of σ_z can be chosen as the one that minimizes the NFA.

In the rest of the paper, we suppose that the precision σ_z was chosen following this remark.

Similarly, the region R and the plane π can be chosen as the ones minimizing the NFA:

Definition 3 *The NFA of a group $G \subset \mathcal{D}$ is defined:*

$$NFA(G) = \min_{\substack{R\in\mathcal{R}, R\supset G \\ \pi\in\Pi_R}} NFA(G, R, \pi) \tag{2}$$

where Π_R is the set of all planes defined for region R.

Remark 2 *In Definition 3, The region and plane that minimize the NFA are respectively the smallest region of $\mathcal{R}$ containing G and the plane that best fits group G. Therefore, a given group can be associated to a plane π and a region R using the NFA.*

Definition 4 (ε-meaningful plane) *A group G is said to be an ε-meaningful planar patch whenever $NFA(G) < \varepsilon$.*

Proposition 1 *Let $\mathcal{S}$ be the set of all the possible pairs (R, π). If we consider a random data set following the background model, then the expected number of ε-meaningful planes in $\mathcal{S}$ is less than ε.*

Proposition 1 is of capital importance in *a contrario* methods. It says that considering the definition of the NFA, less than ε detections of a meaningful plane should happen with a random set. The common use is then to set ε to 1 which means that we should have less than one false detection due to noise. The Proof of proposition 1 is given in appendix A.

Number of tests The number of tests is given by counting all possible region-plane configurations:

$$N_{tests} = \sum_{R\in\mathcal{R}} \#\Pi_R \tag{3}$$

The choice of the set of regions $\mathcal{R}$ is very important. On one hand, a too small set may penalize some groups (*e.g.* $\mathcal{R} = \{\mathcal{D}\}$ is unfair to small groups). On the other hand, a too large set makes it difficult for a group G to be meaningful.

A simple choice for $\mathcal{R}$ is the set of all the rectangles in $\mathcal{D}$ (oriented along the 2 main directions of the map). This choice is similar to the one made in [3] for their *a contrario* clustering method. This set can be reduced without loss of precision by limiting it to rectangles with dyadic size ($2 \times 2, 2 \times 4, 4 \times 4$, etc.).

For the set of planes of a region R, we propose to use the set given by all the triplets of points in R. The number of tests then becomes:

$$N_{tests} = \sum_{R\in\mathcal{R}} \#R \cdot (\#R - 1) \cdot (\#R - 2) \tag{4}$$

Probability of false alarms. Computing the probability $\mathbb{P}[\mathcal{K} \geqslant k]$ depends on the probability of a random point from the background process to be near a given plane. If the precision σ_z is the same for each point, then for any plane π and $\forall X \in R$, we have:

$$\mathbb{P}[\,|\mathcal{Z}(X) - z_\pi(X)| < \sigma_z\,] \leq 2 \cdot \frac{\sigma_z}{z_{max} - z_{min}} = p$$

In such case, $\mathbb{P}[\mathcal{K} \geqslant k]$ can then be upper bounded by the tail of the binomial law:

$$\mathbb{P}[\mathcal{K} \geqslant k] \leq \mathcal{B}(\#R, k, p) = \sum_{j \geqslant k}^{\#R} \left(\begin{array}{c} \#R \\ j \end{array} \right) p^j (1-p)^{\#R-j}$$

When the precision is not constant, $\mathbb{P}[\mathcal{K} \geqslant k]$ can be accurately approximated using the Hoeffding theorem [8]:

$$\mathbb{P}[\mathcal{K} \geqslant k] \leqslant e^{\#R\omega(\eta-\mu,\mu)}$$

with,

$$\omega(\eta,\mu) = (\mu+\eta)\log(\frac{\mu}{\mu+\eta}) + (1-\mu-\eta)\log(\frac{1-\mu}{1-\mu-\eta})$$

and

$$\eta = \frac{k}{\#R}, \quad \mu = \mathbb{E}[\frac{\mathcal{K}}{\#R}] = \frac{1}{\#R}\sum_{X \in R} 2\frac{\sigma_z(X)}{z_{max} - z_{min}}$$

2.3. *A contrario* model selection

Due to its generic form, the NFA can be easily adapted to test the validity of a configuration with several groups, planes and regions. This will allow to choose between two configurations: a simple one with one plane and a more complicated one with several planes. We give here an example for the case of 2 groups.

Following the same reasoning as in Remark 2 to choose regions and planes, and using similar notations as before, we can define the joint NFA for two groups G_1 and G_2 as:

Definition 5 (joint NFA)

$$NFA(G_1, G_2) \equiv N''_{tests} \cdot \mathbb{P}[\mathcal{K}_1 + \mathcal{K}_2 \geqslant k_1 + k2) \quad (5)$$

Note that the number of tests is not the same as before since pairs of regions and planes are now tested:

$$N''_{tests} = \sum_{\substack{R_1 \in \mathcal{R} \\ R_2 \in \mathcal{R} \setminus R_1}} \#\Pi_{R_1} \cdot \#\Pi_{R_2} \quad (6)$$

To be able to do a proper comparison with a single group, the NFA and the joint NFA have to be computed under the same conditions. This means that $G = G_1 \bigcup G_2$, $R = R_1 \bigcup R_2$, and σ_z should be the same for the NFA and the joint NFA. In all that follows, we note $NFA'(G, R_1, R_2)$ the NFA that takes into account two regions for the background process. Note that in such case, the number of tests for the NFA' is changed to:

$$N'_{tests} = \sum_{\substack{R_1 \in \mathcal{R} \\ R_2 \in \mathcal{R} \setminus R_1}} \#\Pi_{R_1 \cup R_2} \quad (7)$$

The smallest value between NFA' and joint NFA determines the best configuration.

3. Plane search

In this section, we propose a heuristic based on the NFA to find planes in disparity maps. Note that other heuristics can be used in conjunction with the NFA as a validation criterion. The algorithm we propose is two-step: first, a topdown dyadic division is done until a good solution is reached, then the resulting groups are merged to refine the result. In both the division and the merging part, the NFA is used as a decision criterion. The main advantage of this method over [10, 7] is to rely only on the 3D information of the points which avoids errors due to image segmentation.

3.1. Splitting step

Starting from a group containing all the points, we divide it into two groups. The best configuration is then kept. If the best configuration is the first group then stop. If it is the division, then for each new group, we start a new division and so on until a good configuration is reached.

To make the division, we chose to use an EM algorithm [5], in order to find the best two normal distributions in the maximum likelihood sense. The choice of normal distributions was made because a planar patch may also be seen as a normal distribution with a low variance according to the plane orthogonal direction.

The decision between one group or two groups is made by following Subsection 2.3. Moreover, we chose to force the division in the case where the main group is not meaningful. To sum up, the one-grouped configuration is kept when:

$$\begin{cases} NFA(G) < 1 \\ NFA'(G, R_1, R_2) < NFA(G_1, G_2) \end{cases} \quad (8)$$

A summary of the whole splitting step is given by Algorithm 1.

3.2. Merging step

Due to the dyadic division process, the final partition of the points might not be optimal. Better configurations may then be found by merging some of the groups. This merging is done only on pairs of adjacent groups, because we are looking for connected closed planes.

Algorithm 1: Splitting step

Data:
G_0, a group of points of $\mathcal{D}$
D_{max}, the maximal depth of the tree
Result:
$\mathcal{G}_{final} = \{G_1, \dots, G_N\}$ such that $G_0 = \bigcup_{i=1}^{N} G_i$

begin
  $\mathcal{G} = \{G_0\}$ is the set of all the groups
  $\mathcal{G}' = \emptyset$ the next set of groups to be tested
  $\mathcal{G}_{final} = \emptyset$ is the set of validated groups
  $D = 0$, the current depth of the tree
  while $D < D_{max}$ **do**
    foreach $G \in \mathcal{G}$ **do**
      $(G_1, G_2) \longleftarrow$ EM algorithm(G)
      $nfa \longleftarrow NFA(G)$
      $nfa1 \longleftarrow NFA(G, R_1, R_2)$
      $nfa2 \longleftarrow NFA(G_1, G_2)$
      if $\big(nfa < 1\big)\&\&\big(nfa1 < nfa2\big)$ **then**
        add G to $\mathcal{G}_{final}$
      else
        Add (G_1, G_2) to $\mathcal{G}'$
      end
    end
    $D = D + 1$
    if $D < D_{max}$ **then**
      $\mathcal{G} = \mathcal{G}', \mathcal{G}' = \emptyset$
    else
      $\mathcal{G}_{final} = \mathcal{G}_{final} \bigcup \mathcal{G}'$
    end
  end
end

Algorithm 2: Merging step

Data:
$\mathcal{G} = \{G_1, \dots, G_N\}$, where $G_i \subset \mathcal{D}$
Result:
$\mathcal{G}' = \{G'_1, \dots, G'_M\}$ such that $\bigcup_{G' \in \mathcal{G}'} G' = \bigcup_{G \in \mathcal{G}} G$

begin
  $\mathcal{G}' = \mathcal{G}$.
  $\mathcal{Q} = \emptyset$ is the priority queue of merges.
  foreach $(G_1, G_2) \in \mathcal{G}^2$ **do**
    if $G_1 \cap G_2$ *are neighbors* **then**
      add (G_1, G_2) to $\mathcal{Q}$
    end
  end
  while $\#\mathcal{Q} > 0$ **do**
    $(G_1, G_2) = \arg\min_{G,G'} F(G, G')$
    Remove (G_1, G_2) from $\mathcal{Q}$
    if $\big(F(G_1, G_2) < 1\big)\&\&\big(NFA(G_1 \cup G_2) < 1\big)$
    then
      The groups are merged
      Remove G_1 and G_2 from $\mathcal{G}'$
      Add $G_{12} = G_1 \cup G_2$ to $\mathcal{G}'$
      foreach k *such that* $(G_1, G_k) \in \mathcal{Q}$ *or* $(G_2, G_k) \in \mathcal{Q}$ **do**
        Remove (G_1, G_k) and (G_2, G_k) from $\mathcal{Q}$
        Add (G_{12}, G_k) to $\mathcal{Q}$
      end
    end
  end
end

The decision of merging or not is made with exactly the same criterion as for the splitting. However, since two groups may be first merged toghether or with a third one, a merge order has to be set.

We introduce a contrast factor to decide the merge order. This contrast factor is defined as:

$$F(G_1, G_2) = \frac{NFA'(G_{12})}{NFA(G_1, G_2)} \tag{9}$$

The lower this contrast factor is the lower the NFA of a single group is compared to the one of two groups. In other words, the lower the contrast factor is the more likely it is to have a configuration using a single group. A priority queue is built using growing contrast factor (the lowest contrast factor is the first).

Whenever groups G_1 and G_2 are merged, each pair of groups including either G_1 or G_2 is removed from the priority queue, and in exchange the corresponding pair formed with the merged group $G_{12} = G_1 \cup G_2$ are added to the priority queue. The complete merging process is shown in algorithm 2. Its effectiveness is illustrated in Figure 1.

3.3. Possible improvements

When the number of planes is too large in a disparity map, dividing groups dyadically might be intensive or might stop in a local minimum. A solution to that is to roughly divide the disparity map into blocks before the splitting process. The splitting step then becomes a lot easier. Moreover, the resulting oversegmentation can be easily overcome by the merging step (see for instance Figure 1).

Another possible improvement is to add image information to this algorithm. For instance, one could use the line segments of the reference image (by using for instance [21]). Indeed, the line segments might suggest separations between different objects in the image. Therefore they might be used to guide the divisions or to try other divisions than the ones given by the EM algorithm.

4. Experimental Results

We tried our algorithm on some of the disparity maps from the Middlebury database [16] with a constant precision parameter of 1 pixel (which is the quantification step of the disparity maps). The results are given in Figure 2 and

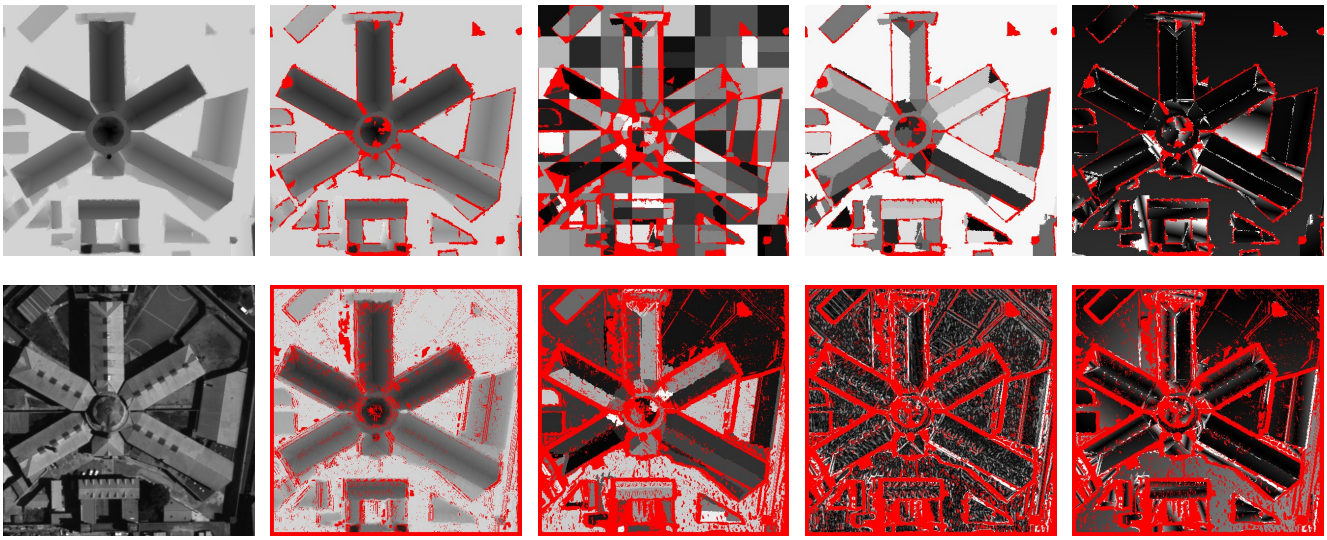


Figure 1. Toulouse's St-Michel Jail (the red parts are unknown or non-planar). *Top row: Planar patch detection on noisy ground truth disparity.* From left to right: ground truth, result of our algorithm after projection on planes, plane classification after splitting, plane classification after merging, absolute value of the difference between the ground truth and our result (black=0 pixels, white = 0.017 pixels). *Bottom row: Planar patch detection on semi-dense stereo correlation results.* From left to right: reference image (secondary image is simulated as a deformation of this image with the ground truth disparity), result [15] (input of our algorithm), plane classification after split & merge algorithm, absolute value of the difference between the ground truth and result of [15], absolute value of the difference between the ground truth and our plane projection for [15] (black=0 pixels, white = 0.075 pixels).

Table 4. The disparity maps of these data sets are sometimes only made of planes (Venus, Sawtooth) and sometimes made of more complex structures. Extremely irregular structures are rejected. If, however, structures are smooth enough, they can be locally approximated by planes up to the given precision and this is the answer of our algorithm. Note that further extensions of our model selection criterion are possible, which should distinguish for instance quadrics from planes.

The piecewise planar approximation is not too simplistic as shown by the error maps given in Figure 2 (e) and the error mesurements of Table 4. The remaining error after projection on the various planes seems to be mostly due to the quantification step of the disparity maps. This explains the oscillations of the errors. Each different period of oscillation corresponds to a different plane. Had the planes been badly estimated, or some of the points been associated to a wrong plane, more errors would be visible. On the another hand, the obtained classifications do not seem to be oversegmented. The various periods of oscillations seem to correspond to one plane most of the time.

Our approach is robust to change of precisions and gives a planar approximation at different scales (see Figure 4). This is usually not the case of RANSAC based approaches which tend to fail when the precision is not optimal.

We also tried our algorithm on a simulated stereo pair of Toulouse's St-Michel Jail. For this experiment, a subpixel precision is expected. Two tests were performed: *(i)* We added an additional Gaussian noise of variance 0.02 pixels to the ground truth disparity. The results are given in Figure 1 and Table 4, column Toulouse. For this experiment, most of the planes seem to have been detected. The error was significantly reduced (Figure 1 (f)) and is mostly localized around the edges (RMSE = 0.007 with edges). However, the planes were estimated with a great precision (RMSE = 0.005 pixels without the edges).

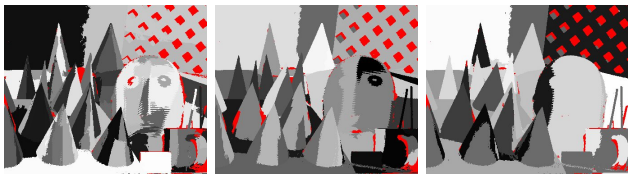
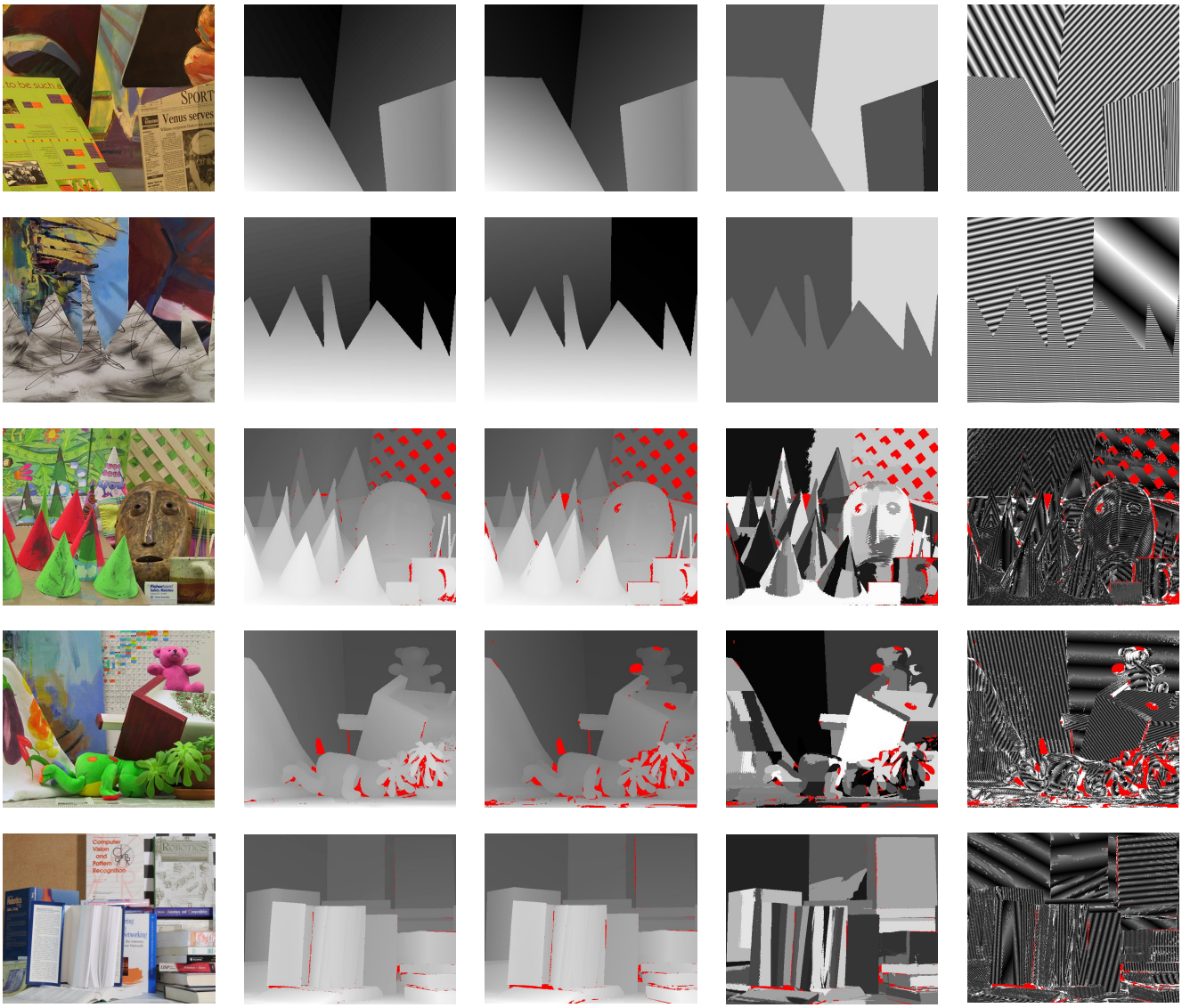


Figure 3. Classification obtained with different precisions. From left to right: precision=1 pixel, 85 planes detected; precision=2 pixels, 44 planes detected; precision=5 pixels, 23 planes detected.

The algorithm in [15] was used on a simulated low-baseline stereo pair to obtain a disparity map with precision 0.024 pixels. After grouping with our algorithm and reprojection of the measured disparity data on the detected planes we observe that the RMSE is reduced to 0.021 pixels (see Table 1, column Toulouse2) which proves the correctness of our planar approximation.

5. Conclusion

We presented an algorithm for optimally grouping 3D point clouds into planar patches. Inspired from computa-



(a) Reference image (b) Ground truth (c) Projection on planes (d) Planar classification (e) $|(b)-(c)|$

Figure 2. Results obtained with Middlebury's ground truth datasets. From top to bottom, Venus, Sawtooth, Cones, Teddy and Books.

	mixed		planar		non-planar		
	Toulouse	Toulouse2	Venus	Sawtooth	Cones	Teddy	Books
Initial variance(σ)	0.02	0.024	1	1	1	1	1
RMSE	0.005	0.021	0.29	0.29	0.95	1.15	1.01
Error$\geqslant \sigma$ (%)	1.8	-	0	0	3.9	6	3.4
Error$\geqslant 2\sigma$ (%)	0.7	-	0	0	0.8	1.9	0.9

Table 1. Error measurements. First line initial: precision of the ground truth. Second line: RMSE (ℓ_2 error). Third and fourth lines: percentage of outliers.

tional gestalt theory [6], it allows the use of simple grouping laws to robustly detect simple patterns (planar patches here), and to apply later these laws recursively (for instance symmetry or similarity of planar patches) in order to obtain more complex structures, like those proposed in Lafarge's dictionary [13], without making an a priori explicit list of all possibilities. As opposed to the method proposed in [10, 7], our algorithm do not rely on an initial segmentation which can be error prone. The various parameters can be easily set which makes it almost automatic. The ε value can be set to 1 due to its statistical meaning and the precision over the points can be either computed with a method similar to the one of [15] or estimated as the one minimizing the NFA.

Our experiments show that the proposed approach is ca-

pable of detecting a reasonable piecewise affine decomposition even in complex scenes (as opposed to RANSAC based approaches). Moreover, the corresponding regularization reduces the error of punctual disparity measures.

Several applications and improvements are envisaged. The piecewise planar grouping can be used as a basis for interpolation and vectorization algorithms. However, these applications will require a stronger use of luminance (as a post processing refinement of the boundaries between several planes).

6. Aknowledgments

The authors acknowledge support by CNES, UdelaR, the ANR Callisto, and ECOS Sud U06E01 projects. The St-Michel dataset was kindly provided by CNES, its copyright holder.

A. Proof of Proposition 1

Let's note S the random variable defined as:

$$S = \sum_{(R,\pi)\in\mathcal{S}} \chi_{(R,\pi)}$$

where, $\chi_{(R,\pi)} = \mathbb{1}_{(R,\pi) \text{ is } \epsilon\text{-meaningful}}$. Using the linearity of the expectation operator one has:

$$E[S] = \sum_{(R,\pi)\in\mathcal{S}} E[\chi_{(R,\pi)}] \tag{10}$$

Then, using definition 2 we can write: $E[\chi_{(R,\pi)}] = P(\mathcal{K} \geqslant k(\varepsilon)) \leqslant \frac{\varepsilon}{N_{tests}}$. At last using the definition of the number of tests given in 2.2, and substituting it in equation 10, we obtain the result:

$$E[S] \leqslant \sum_{(R,\pi)\in\mathcal{S}} \frac{\varepsilon}{N_{tests}} = N_{tests} \cdot \frac{\varepsilon}{N_{tests}} = \varepsilon \tag{11}$$

References

[1] H. Akaike. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19:716–723, 1974. 2

[2] C. Baillard and A. Zisserman. Automatic reconstruction of piecewise planar models from multiple views. In *CVPR, 1999. IEEE Computer Society Conference on.*, volume 2, page 565 Vol. 2, 1999. 2

[3] F. Cao, J. Delon, A. Desolneux, P. Muse, and F. Sur. A unified framework for detecting groups and application to shape recognition. *Journal of Mathematical Imaging and Vision*, 27(2):91–119, February 2007. 3

[4] J. Delon and B. Rougé. Small baseline stereovision. *Journal of Mathematical Imaging and Vision*, 28(4):209–223, July 2007. 1

[5] A. Dempster, N. Laird, and D. Rubin. Maximum-likelyhood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society Series B*, 39, 1977. 4

[6] A. Desolneux, L. Moisan, and J.-M. Morel. *From Gestalt Theory to Image Analysis: A Probabilistic Approach.* Springer Verlag, 2008. 2, 7

[7] G. Facciolo and V. Caselles. Geodesic neighborhoods for piecewise affine interpolation of sparse data. In *ICIP 09*, 2009. 2, 4, 7

[8] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58:13–30, 1963. 4

[9] A. Hoover, G. Jean-Baptiste, X. Jiang, and P. Flynn. A comparison of range image segmentation algorithms. *IEEE Transactions PAMI*, 18(7):673–689, 1996. 2

[10] Igual. Automatic low baseline stereo in urban areas. *Inverse Problems and Imaging*, 1(2):319–348, May 2007. 2, 4, 7

[11] X. Jiang and H. Bunke. Fast segmentation of range images into planar regions by scan line grouping. *Machine Vision and Applications*, 1994. 2

[12] P. Labatut, J.-P. Pons, and R. Keriven. Hierarchical shape-based surface reconstruction for dense multi-view stereo. *The 2009 IEEE International Workshop on 3-D Digital Imaging and Modeling*, 09. 2

[13] F. Lafarge, X. Descombes, J. Zerubia, and M. Pierrot-Deseilligny. Automatic building extraction from dems using an object approach and application to the 3d-city modeling. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(3):365–381, May 2008. 2, 7

[14] J. Rabin, J. Delon, Y. Gousseau, and L. Moisan. Macransac: Détection automatique d'objets multiples. *RFIA'10*, 10. 2

[15] N. Sabater. *Reliability and Accuracy in Stereovision Application to Aerial and Satellite High Resolution Images*. PhD thesis, ENS Cachan, 12 09. 1, 2, 3, 6, 7

[16] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *IJCV*, 47:7–42, April-June 2002. 5

[17] G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 2:461–464, 1978. 2

[18] R. Taylor, M. Savini, and A. Reeves. Fast segmentation of range imagery into planar regions. *Computer Vision, Graphics, and Image Processing*, 45:42–60, 1989. 2

[19] R. Toldo and A. Fusiello. Robust multiple structures estimation with j-linkage. In *ECCV '08*, pages 537–547, Berlin, Heidelberg, 2008. Springer-Verlag. 1

[20] A. Torii and A. Imiya. The randomized-hough-transform-based method for great-circle detection on sphere. *Pattern Recognition Letters*, 28(10):1186–1192, July 2007. 1

[21] R. G. von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall. Lsd: A fast line segment detector with a false detection control. *PAMI'08*, 2008. 5

[22] W. Zhang and J. Kosecká. Nonparametric estimation of multiple structures with outliers. In *ECCV 06*, pages 60–74, 2006. 1

[23] M. Zuliani, C. S. Kenney, and B. S. Manjunath. The multi-ransac algorithm and its application to detect planar homographies. In *ICIP*, 2005. 1