1

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Robust matching by partial correlation

Lan, Zhong-Dan / Mohr, Roger

N RR-2643

Août 1995

PROGRAMME 4

Robotique,

image

et vision

Rapport de recherche

1995

ISSN 0249-6399

Robust matching by partial correlation

Lan, Zhong-Dan / Mohr, Roger

Programme 4 — Robotique, image et vision

Projet MOVI (common to CNRS, INRIA, INPG and UJF)

Rapport de recherche n° RR-2643 — Août 1995 — 21 pages

Abstract: Visual correspondence problem (matching), is an major issue in vision. Matching can be divided in *feature-based* and *area-based* matching. In area-basedmatching, correlation is a common tool for the visual correspondence problem. But most classical correlation methods fail near the disparity discontinuity, which occurs at the boundaries of objects. In this report, we propose a *partial correlation* technique for solving this problem. We use a robust statisticstool to find the *good part* to be considered for applying correlation, We have compared our method with the standard and non standard techniques. Experimental results validate the approach.

Key-words: vision, matching, correlation, robustness, occlusion.

(Résumé : tsvp)

*

Unité de recherche INRIA Rhône-Alpes
46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1 (France)
Téléphone : (33) 76 57 47 77 – Télécopie : (33) 76 57 47 54

Appariement robust par corrélation partielle

Résumé : L'appariement entre deux images est une clé de nombreux problèmes de perception, en particulier celui de la perception stéréoscopique. Ce problème est essentiellement abordé par deux types d'approches, celle basé sur l'utilisation des caractéristiques comme les contours, et celle basé sur la mise en correspondance d'une partie du signal. La corrélation est l'outil le plus commun pour la second catégorie. Mais les approches classiques de la corrélation ont des difficultées près des discontinuités qui se présentent à la limite des objets occultants. Dans ce rapport, nous proposons une technique baptisée *corrélation partielle* pour résoudre ce problème. Nous utilisons un outil de la statistique robuste pour trouver la —*bonne partie* des fenêtres à corréler. Nous avons comparé notre méthode avec les méthodes standards et les méthodes plus robustes. Des expérimentations valident la méthode.

Mots-clé : vision, appariement, corrélation, robustesse, occultation.

Contents

1 Introduction

The visual correspondence problem, i.e. matching between two images is an major task in vision [14], [3], [10], as it underlies the computation of motion (motion matching) and the computation of stereo depth (stereo matching). Given two images of a same scene, a pixel in one image corresponds to a pixel in the other one if they are projections along lines of sight of the same physical scene element. If they are temporally consecutive, then the computation of correspondence determines motion. If they are spatially separated but simultaneous, then the computation of correspondence determines stereo disparity.

Matching can be divided in area based matching and feature based matching (see for instance [8] and [2] in the case of stereo-vision). Three types of methods have been used for area based matching:

- optical flow [9]
- Fourier transformation [22]
- correlation [1]

In this report, we investigate only the third method. People interested in the two previous ones might have a look in the referenced papers. The use of correlation as a similarity measure between two signals is a well known technique. It is commonly used in stereo vision for the visual correspondence problem [5] and extensive comparison have been made for the evaluation of different correlation criteria [1].

As indicated in [16], [6] and [12], occlusion plays an important role in stereo matching. In [12] Intille uses Ground Control Points (GCP) to construct the best disparity path; this method gets its best results if GCP are effectively very close to each occluding boundary as this is the region in which all methods fail.

We came to the conclusion that only the use of a *robust methods* can overcome the matching problems every method encounters at this occluding contours. This report is devoted to describe a method which provides good matches in these areas, providing therefore a tool for finding good GCP points where they are really needed. For this purpose, a new correlation algorithm is developed which works even when the occlusion arises. The key idea is to use a robust estimator to find the *good part* on which the correlation should be performed, and then to carry out the *partial* correlation on this part.

Robust estimator were already used for solving matching problems. For example, in the paper of Odobez and Bouthemy [18], a *robust* optical-flow method basing on `M`-estimators is used to establish the motion correspondence. Zabih [23] uses a *non parametric method* to compute the visual correspondence for stereo-vision. We shall come back to the Zabih's correlation methods and show some comparison between his methods, ours and more standard methods.

The paper is organized as following. Firstly we introduce some related work, then we explain our model for the *partial correlation* and introduce the robust estimators developed in *robust statistics* [11] [7], especially the `LMedS`-estimator, which is used in our algorithm to find the *good* part of the correlation on which the *partial correlation* should be done. At the end we show experimental results and propose some future work.

2 Related work

2.1 Correlation and image transforms

Visual correspondance of two vectors is usually defined by correlation. There are many different definitions for this purpose (see the Table 1 in [1] for instance). Experimental comparison were perforemed and are reported in [1] and in [5]. Among them, we have chosen the zero mean sum of squared differences because of its relative simple form and its good experimental performance to varying lighting conditions:

$$\mathrm{ZSSD}(X, X+dX) = \sqrt{\frac{\sum_{\Delta\in W}((I_1(X+\Delta) - \bar{I}_1(X)) - (I_2(X+\Delta+dX) - \bar{I}_2(X+dX)))^2}{(2N+1)(2M+1)}}$$

where $X = (x, y)$, $\Delta = (u, v)$, W is the window defined by $\{(u,v) | -N \leq u \leq N \text{ and } -M \leq v \leq M\}$, $dX = (dx, dy)$ is the disparity and $\bar{I}_i(X)$ is the mean gray level value of image i in the window $(i = 1, 2)$.

Alternative approaches have been proposed which suggest to use a transformed version of the original signal. A standard transform is to use the Laplacian of the image. In some way this is equivalent to normalizing the image as it is done by the zssd method.

However these methods are strongly corrupted by local strong noise or by the presence of local perturbation of the original signal when occlusion arises. In order to reduce such a *factionalism* authors limited the influence of the numerical values by taking into account only the sign of the Laplacian [24] [17], or the orientation of the gradient [21]. In such a case, the change introduced in the correlation answer by local perturbations will be in direct ratio with the size of the perturbed area. A similar idea is exploited in the rank and census transform presented in the next section.

2.2 Rank and census transform [23]

Zabih has developed an approach which relies on local transforms based on non-parametric measures. They are designed to tolerate *factionalism*. Non parametric statistics [13] is distinguished by the use of ordering information among data, rather than the data values themselves.

Two local non-parametric local transforms were introduced. The first one, called *rank transform*, is a non-parametric measure of local intensity. And the second one, called *census transform*, is a non-parametric summary of local spatial structure.

Let P be a pixel, $I(P)$ its intensity, and $N(P)$ the set of pixels in its neighbour. All non parametric transforms depend upon the comparison of $I(P)$ and the intensity of the pixels in $N(P)$. Define $\xi(P, P')$ to be 1 if $I(P') < I(P)$ and 0 otherwise. The non-parametric local transforms depend solely on the set of pixel comparisons, which is the set of ordered pairs :

$$\Xi = \bigcup_{P' \in N(P)} (P', \xi(P, P'))$$

The *rank* transform is defined as the number of pixels in the local region whose intensity is less than the intensity of the center pixel. Formally, the rank transform $R(P)$ is

$$R(P) = ||\{P' \in N(P) | I(P') < I(P)\}||$$

For the rank transform, the L_1 correlation (minimization of the sum of absolute value of differences) on the transformed image is used on order to preserve a response that diverges linearly with the number of outliers.

The *census* transform $R_\tau(P)$ is a mapping from the local neighbour surrounding a pixel P to a bit string representing the set of neighboring pixels whose intensity is less than that of P. Let $N(P) = P \oplus D$ where $\oplus$ is the Minkowski sum and D is a set of displacements, and let $\otimes$ denote concatenation. The census transform can then be specified,

$$R_\tau(P) = \bigotimes_{[i,j] \in D} \xi(P, P + [i, j])$$

Two pixels of census transformed image are computed for similarity using the Hamming distance, i.e. the number of bits that differ in two bit strings. The algorithm computes the correspondence by minimizing the Hamming distance after applying the census transform.

2.3 A modification of Zabih's approach

And based on his idea, we have developed a local transform which captures more information and therefore becomes more complex. Instead of classifying $(I(P) < I(P'))$ by 0 or 1, we replace the binary value by a probability of $(I(P) < I(P'))$, more precisely:

given a pixel P and a neighboring pixel $P \oplus [i, j]$, the greater the $I(P) - I(P \oplus [i, j])$, the more confidence we have that $I(P \oplus [i, j]) < I(P)$. Of course, the confidence depends also on the standard deviation of $I(P) - I(P \oplus [i, j])$.

Let the probability of the $I(P \oplus [i,j]) < I(P)$ to be $G(s/\sigma)$ where s is the observed value of $I(P) - I(P \oplus [i,j])$, σ is the estimated standard deviation of $I(P) - I(P \oplus [i,j])$, and G is the Gaussian distribution function . So the local transform depends also upon the estimation of the standard deviation of $I(P) - I(P \oplus [i,j])$. It was estimated in the following way:

- estimate σ_1 = the standard deviation of $I(P)$
- estimate σ_2 = the standard deviation of $I(P \oplus [i,j])$
- estimate r = their correlation coefficient
- the standard deviation of $I(P) - I(P \oplus [i,j])$ can now be computed as $\sqrt{\sigma_1^2 + \sigma_2^2 - 2r\sigma_1\sigma_2}$

From our observation, we have put the σ_1 and σ_2 to be constant 5, r to be $\rho_h^i * \rho_v^j$, where we put $\rho_h = \rho_v = 0.9$.

Of course this method is more flexible than the Zabih's one, but it cannot be encoded as efficiently as it was done for the computation of the Hamming distance. From our experiments, we estimated that the computation time for this method is about 100 times larger than the finely tuned method developed by Zabih.

3 Partial correlation

Let's select in the first image a template of fixed size. Suppose that there are several candidates of the same size. Certainly, we can use the *epipolar constraint* and other constraint to reduce the amount of the candidates. What we used after are the *region of interest* constraint and the *epipolar constraint*. These two constraints are displayed in the Figure 1: the window candidate has to be centered on the corresponding epipolar line and stay within the region of interest.

3.1 Our partial correlation model

Most correlation techniques have difficulty near disparity discontinuities, or in places where highlights occur, as the window under consideration is locally partially but hardly corrupted in small area. Therefore in such a case the pixels in a local region represent scene elements from two distinct intensity populations. Some of the pixels correspond to the template under consideration and some from another parts of the scene. As we already mentioned, this leads to a problem for many correlation techniques, since the standard correlation techniques are usually based on standard statistics methods, which are best suited to a single population. We shall call this phenomenon *partial occlusion* even it might have other physical sources as highlights, and we propose the *partial correlation* idea to overcome it.

Consider the two windows of the Figure 2, they match except for the right-up part in the image I_1 and I_2. We have to recover this corrupted part. We assume that locally the signal obeys an affine

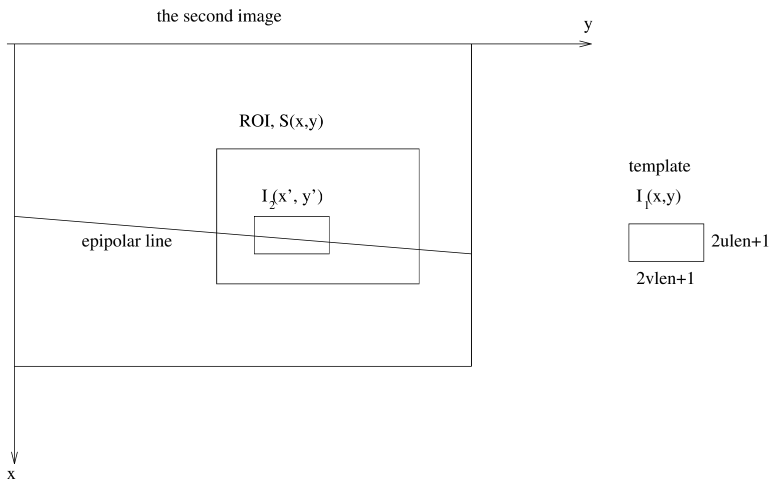

Figure 1: Epipolar constraint and region of interest constraint

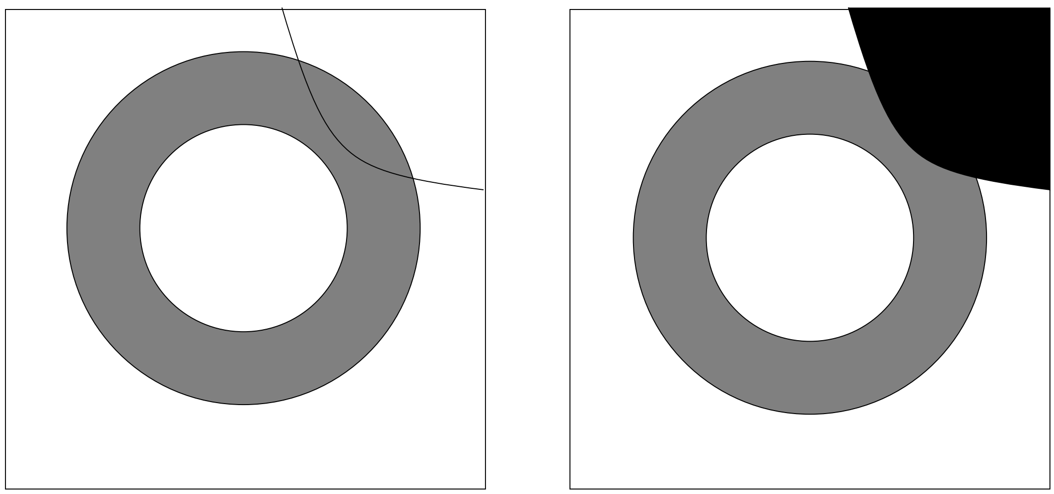

Figure 2: Occlusion.Left: I_1, right: I_2

relation from on image to the other (see Figure 3):

$$I_2(s,t) = kI_1(s+dx, t+dy) + l + \epsilon$$

where (dx, dy) is the disparity of the point (x, y), $s = x+u, t = y+v$ and $s \in \{x - N \ldots x + N\}$ $t \in \{y - M \ldots y + M\}$, $2N + 1$ and $2M + 1$ are the width and the height of the template window respectively and ϵ is a Gaussian image noise.

Such an affine relation is true in the window except for the parts *occluded* (as the right-up part indicated in the Figure).

A simpler model would allow just a shift in intensity, i.e. the scale factor k is set to 1 :

$$I_2(s,t) = I_1(s+dx, t+dy) + l + \epsilon$$

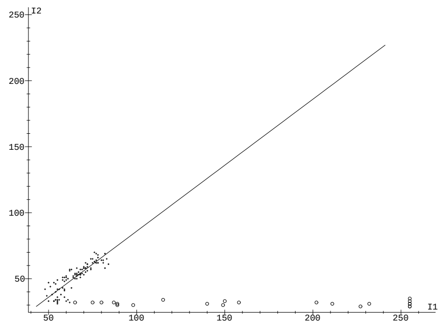

Figure 3: Affine relation between I_1 and I_2, the small circles are outliers

If the center of the window with coordinates (x, y) is *not* occluded, an additional constraint can be stated; it will be called the(*center point constraint*) here after. It just states that the center of the window has to satisfy the previous affine or translational constraint:

$$\begin{aligned} I_2(x,y) &= kI_1(x+dx, y+dy) + l \text{ (affine case)} \\ I_2(x,y) &= I_1(x+dx, y+dy) + l \text{ (translation case)} \end{aligned}$$

The occluded part is found using *the robust statistics* where are detailed here after the section 3.2. Application of this technique shall be illustrated in section 4

Having found the *occluded* parts, the correlation is restricted on the remaining parts of the two windows. We call this technique *partial correlation*. In fact *partial correlation* is a special case of *weighted correlation*, as we leave $\omega(u,v) = 1$ (inlier) or $\omega(u,v) = 0$ (outliers).

For example, the weighted $ZSSD$ can be expressed as follows :

$$ZSSD_w(X, dX) = \sqrt{\frac{\sum_{\Delta \in W}(((I_1(X+\Delta) - \bar{I}_1(X)) - (I_2(X+\Delta) - \bar{I}_2(X)))^2 weight(\Delta))}{\sum_{\Delta \in W} weight(\Delta)}} \tag{1}$$

3.2 Robust estimator for linear regression

The intensity of a window corresponding to a template is defined by one or two parameters depending if we consider the affine or translational transformation of light intensity. Estimating these parameters in presence of outliers and at the same time founding the outliers is described here under the general framework of robust estimation.

3.2.1 The challenge of outliers

As in many reel world situation, outliers corrupt the data and therefore the method of least mean squares fail. This is why we should formulate the algorithms using methods coming from *robust statistics* [20].

We suppose that vision is a processes to fit some visual models to some unreliable data. Here unreliableness means the appearance of some *outliers*. Regression analysis (fitting a model to noisy data) is an important statistics tool in vision. The usually used regression method, called the least mean squares regression is optimum when the noise is Gaussian. However this method is not robust to outliers. See Figure 4 for an example [19]. In this figure, an outlier corrupt totally the result.

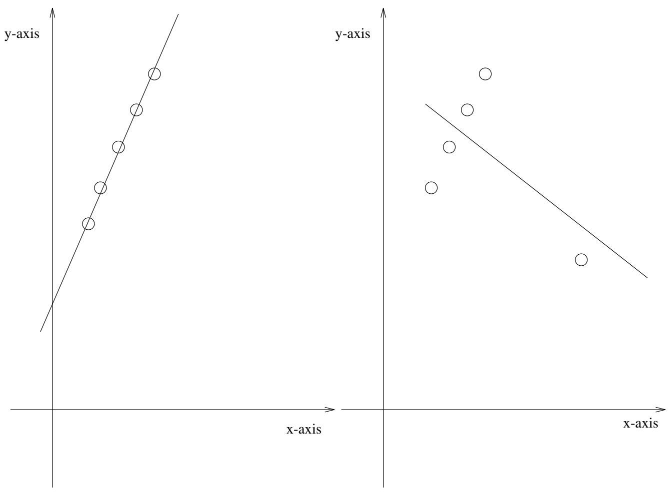

Figure 4: (left) Original data with five points and their line by the least mean squares regression (right) Same data but with an outlier in the direction x

More mathematically, a linear regression model of order n is represented by the equation :

$$\hat{y}_i = \hat{\theta}_1 x_{i1} + \hat{\theta}_2 x_{i2} + \ldots + \hat{\theta}_n x_{in}$$

for the i^{th} point of the data.

$\hat{\theta}_i$ is the estimation of the parameters θ_i of the model. The residual for each point of the data is $r_i = y_i - \hat{y}_i$.

In the least mean squares regression, the parameters' estimation of the model is obtained by the minimization of the sum of the squared residuals :

$$\min \sum_{i=1}^{n} {r_i}^2$$

The *breakdown point* of a regression method, noted ${\epsilon_n}^*$, is the smallest amount of outlier contamination that may force the value of the estimate outside an arbitrary range. See [19] for its more mathematical discussions.

For the least mean squares method, we have ${\epsilon_n}^* = \frac{1}{n}$ and in the limits ${\epsilon_\infty}^* = 0$. That is to say it is not robust at all.

3.2.2 Robust regression

Fortunately, some robust regression methods were developed in robust statistic. Among which, the basic robust estimators are classified as M-estimators, R-estimators and L-estimators [11] [15]. Several robust estimators having the breakdown point close to 0.5 were developed [19]. In this paper, we will use the least median of squares (*LMedS*) regression to find the *outlier* part of the correlation.

In this method, we minimize the median of the squares of the residual :

$$\min_{\hat{\theta}} \operatorname*{med}_{i} {r_i}^2$$

Seeing that the *LMedS* estimator is highly un-analytical, we use the Monte Carlo technique to implement it. See the Table 1 for the algorithm.

This method has the breakdown point 50%. See Fig 5. In this figure, the result is not corrupted by the outlier.

To find the outliers, we use the $\hat{\theta}$ obtained by the *LMedS* estimator and compute every residual r_i. As is described in [15], we can estimate the standard deviation of noise as

$$\sigma_i = 1.4826(1 + \frac{5}{n-p})\sqrt{\operatorname*{med}_{i} {r_i}^2}$$

Assume that there are n data points and p parameters in the linear model.

1. Choose p points at random from the set of n data points
2. Computer the fit of the model to the p points
3. Computer the median of the square of the residuals

The fitting is repeated until a fit is found with sufficiently small median of squared residuals or until some predetermined number of re-sampling steps.

Table 1: The least median squares regression algorithm

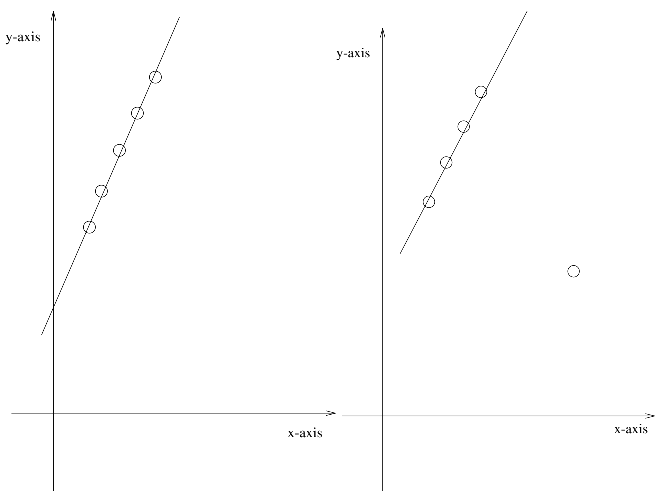

Figure 5: (left) Original data with five points and their line by regression of the least median squares (right) Same data but with an outlier in the direction x

Here n is the data number and p is the parameter number. The factor 1.4826 is for consistent estimation in the presence of Gaussian noise, and the term $\frac{5}{n-p}$ is recommended by Rousseeuw and Leroy [19] as a finite sample correction. Based on the robust LMedS model and the standard deviation estimate binary weights can be allocated to the data points[15]:

$$\omega_i = \begin{cases} 1 & \frac{|r_i|}{\hat{\sigma}} \leq 2.5 \\ 0 & \frac{|r_i|}{\hat{\sigma}} > 2.5 \end{cases}$$

4 Experimental results and discussion

In this section, first some isolated examples are shown in 4.1 and then statistic comparisons are made with standard and non standard methods in 4.2

4.1 Examples

To prove our approach, we've tested it on some points taken from [16] where the classical method fails. In Figure 6, two examples are displayed. The two rectangles in the right image are the region of interest chosen. Standard correlation techniques failed to find the correct corresponding point; in one case a huge occlusion occurs and in the second case the surface is almost tangent to the viewing direction and the windows dont fit correctly.

Epipolar lines are displayed as are displayed the wrong matches found by the standard zssd method. Partial correlation succeed in each of these cases. For instance, on the Paolo's finger, the partial correlation method finds the correct match on the right, where zssd finds the false match on the left.

In the Figure 7, the outlier points are displayed for each template, when the correct match is found.

From these two examples, we see that our method might work where the classical one fails. We are going to quantify this by collecting statistics on large experiments.

4.2 Statistics results

In order to collect the exact result of matching, experiments were conducted on planar patches. In such a case the exact matches can be computed by estimating precisely the homography mapping on patch on the other. We have only tested the method using the translation relation and the center point constraint. We call it rzssdc here, which stands for *robust centered zssd*. rzssdc is compared with the standard zssd, the rank method and our modified implementation of the census transform.

Figure 6: Paolo. Two match tests: one on the finger and the other on the ear. The small squares in the left image are the templates.

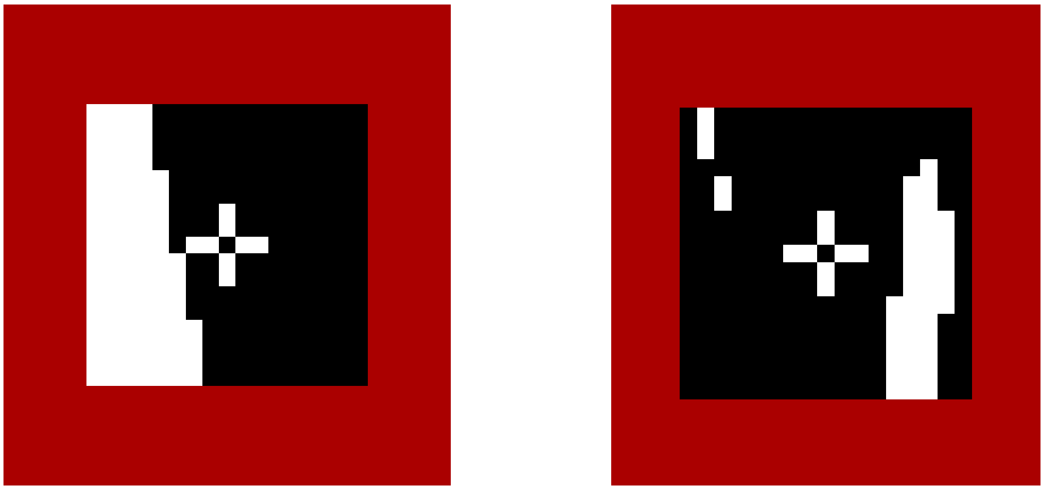

Figure 7: Outlier parts (white) found for the two examples. In the left image is displayed the outlier part for the finger and in the right one is displayed the outlier part for the ear.

Two pairs of images are used for this test. They are displayed in Figure 8 and Figure 9.

Figure 8: Stereo pair 1: *benetton*

Figure 9: Stereo pair 2: *world*

Both of them have a planar background but these backgrounds are set with different kinds of texture. Four tests were designed. Figure 10 displays two sets of points for the *benetton* background for which the corresponding points were sought. Figure 11 displays the test 3 and 4 which were performed with the *world* planar background; for the test 3 the template points are scattered along the occluding contour: a difficult test case.

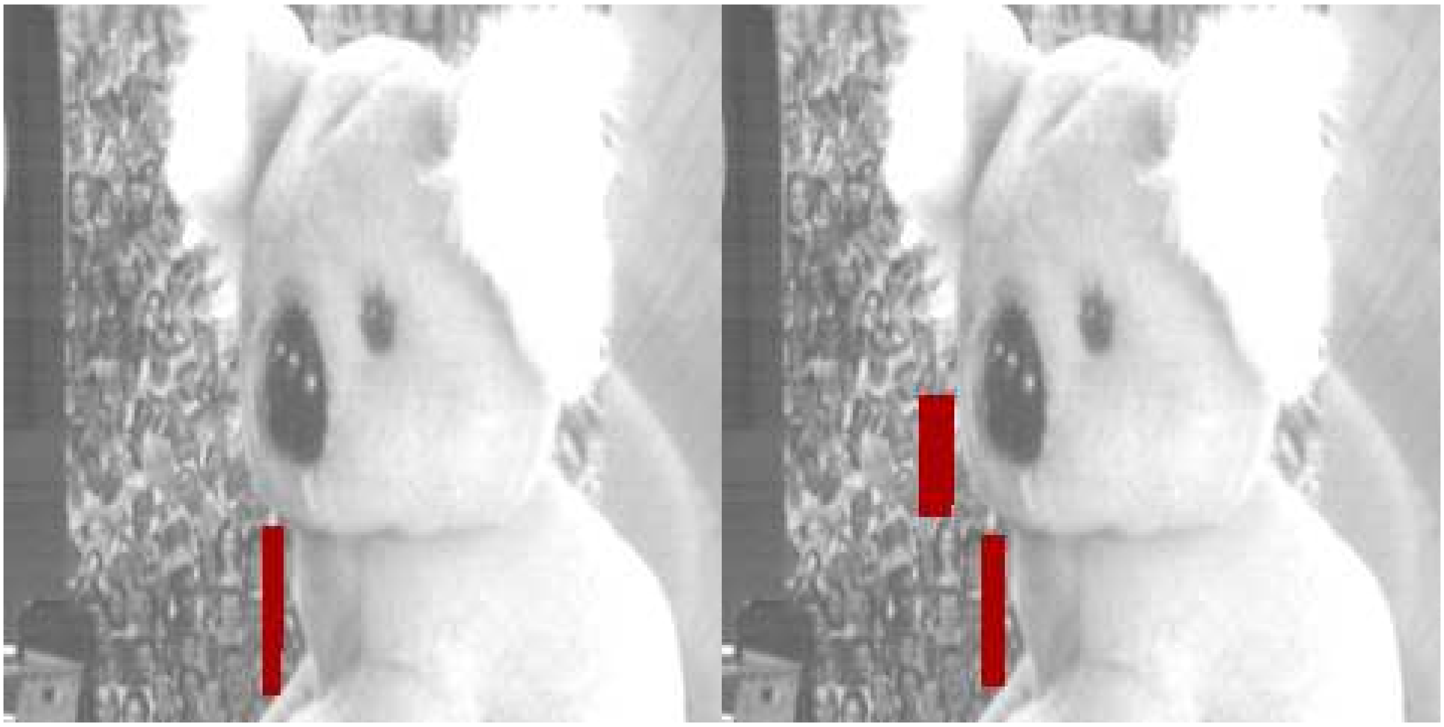

Figure 10: Background *benetton*. Points selected (in black) for test 1 and 2. Left: example1 corresponding to Table 2; Right: example 2 corresponding to Table 3

The different methods were run on these four tests. The candidates were constrained to stay on the epipolar line and to have the disparity less than 200. For each method, we compare the results found with the exact value provided by the homography.

The results are reported in the Table 2, 3, 4 and 5. For each method (i.e. : rank, census, rzssdc and zssd) are indicated:
- the number of accurate matches: up to one pixel error (good match),
- the number of matches in a distance between one and two pixels error (near miss),
- the number of matches at a distance of two to three pixels (bad matches),
- the matches which lie more than three pixels away from their exact position (false matches).

4.3 Discussions

From the previous tables, we see that the rzssdc method provides the best results, with less than 5% wrong matches. Census method gets the second best results but and reaches between 5 and 33% wrong matches; as already mentioned by Zabih, the rank method is even less good. The zssd

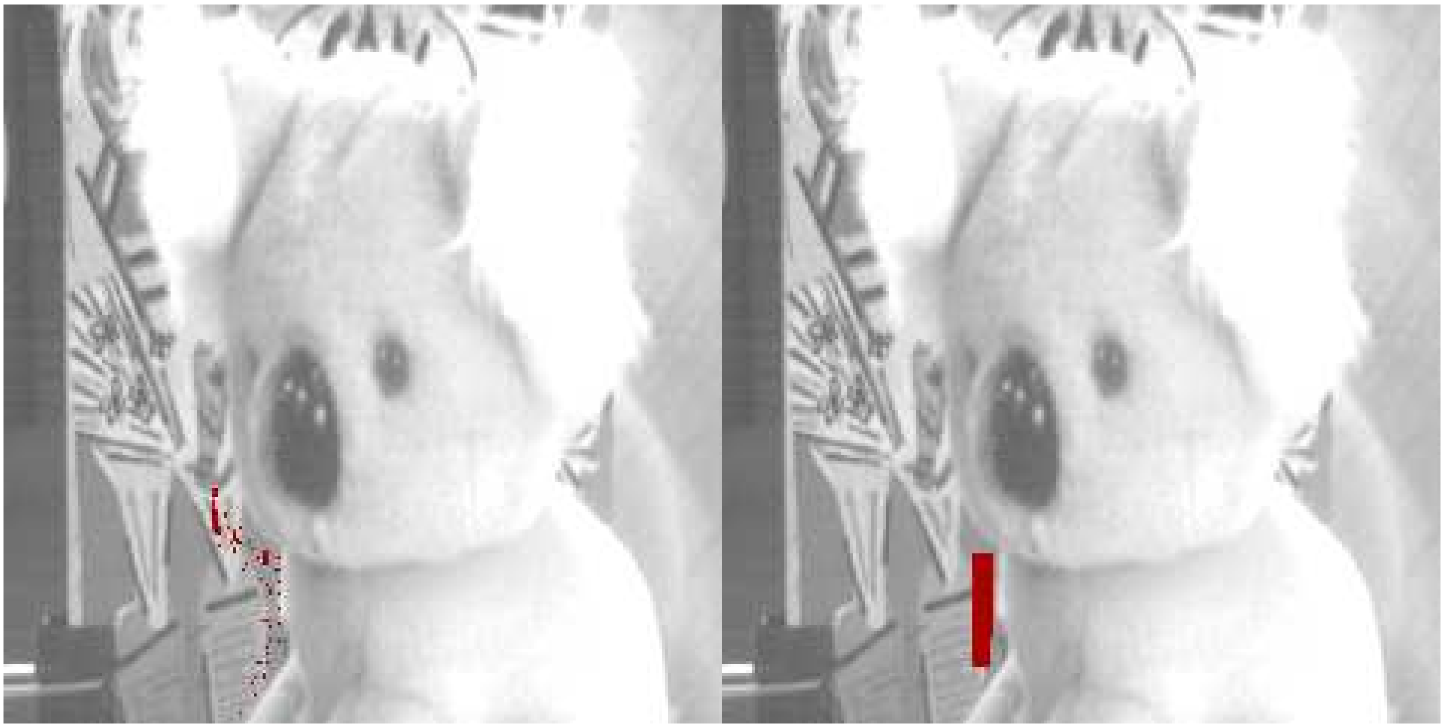

Figure 11: Background *world*. Points selected (in black) for test 3 and 4. Left : example 3 corresponding to Table 4; Right : example 4 corresponding to Table 5

rank			
0–1	1–2	2–3	3–infty
246	96	0	57
census			
0–1	1–2	2–3	3–infty
305	63	0	31
rzssdc			
0–1	1–2	2–3	3–infty
316	63	0	20
zssd			
0–1	1–2	2–3	3–infty
188	37	0	174

Table 2: the first example on the image *benetton*

rank			
0–1	1–2	2–3	3–infty
713	98	0	89
census			
0–1	1–2	2–3	3–infty
773	83	0	44
rzssdc			
0–1	1–2	2–3	3–infty
796	74	0	30
zssd			
0–1	1–2	2–3	3–infty
658	38	3	201

Table 3: the second example on the image *benetton*

rank			
0–1	1–2	2–3	3–infty
77	14	0	49
census			
0–1	1–2	2–3	3–infty
89	12	0	39
rzssdc			
0–1	1–2	2–3	3–infty
118	10	0	12
zssd			
0–1	1–2	2–3	3–infty
56	0	0	84

Table 4: the first example on the image *world*

rank			
0–1	1–2	2–3	3–infty
129	15	0	122
census			
0–1	1–2	2–3	3–infty
185	6	0	75
rzssdc			
0–1	1–2	2–3	3–infty
234	29	2	1
zssd			
0–1	1–2	2–3	3–infty
71	1	0	194

Table 5: the second example on the image *world*

rank			
0–1	1–2	2–3	3–infty
11946	2512	299	968
census			
0–1	1–2	2–3	3–infty
12960	2237	105	423
rzssdc			
0–1	1–2	2–3	3–infty
11845	1985	193	1702
zssd			
0–1	1–2	2–3	3–infty
13140	1941	116	528

Table 6: An example for the general case on the image *world*

is the worse in such difficult conditions and its scores between 22 and 72% of wrong matches. This experience validates the factionalism of the census method, and proves that with an additional computational effort this can improved with the partial correlation method.

However, *if we choose points in non occluded regions, we get the best result using the census or the classical method.* Table 6 provides the results for the four previous method tested on a 125×125 area. Notice that all other methods do not behave too badly too. More than 80% matches are within 2 pixels for each method.

There are several explanations for explaining these last results. Firstly as census and zssd consider the whole window, the support for their answer is larger and more robust to noise. Partial correlation might select a wrong window with perfect partial matches instead of considering the signal globally when it should be done. This is illustrated in Figure 12 on a 1 dimensional signal.

Therefore we think that the *partial correlation* is good to be used only if occlusion is suspected. The fundamental difficulty for this technique is the conflict between the *similarity* and the *completeness* of the correlation :

Given a template, and a several candidates, which one is the most similar to the template ?

Two subparts of windows might match perfectly (*similarity*); the whole windows might match also (*completeness*) but no more perfectly. As these two criterion compete, a syntheses has to be made; this is not quite easy as completeness is not important at occlusion edges, but important in other areas.

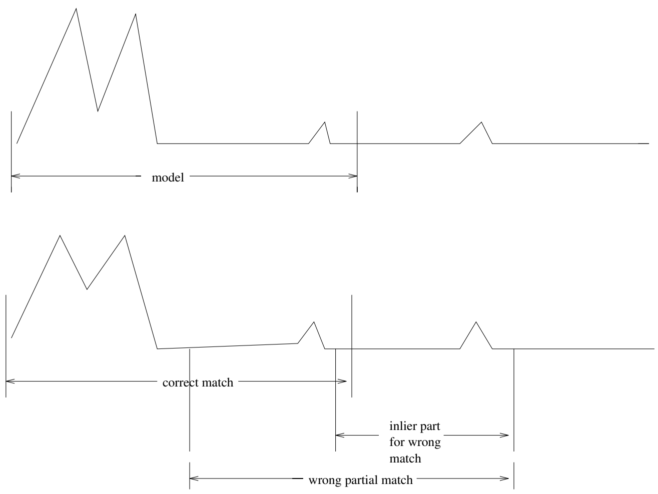

Figure 12: The signal shift, the left part is found to be outlier and rejected, and the right part match very well. If there is no shift, we'll find no outlier part and the whole parts match less well then the shifted one.

5 Conclusion and future work

In this report, we've discussed the matching problem in presence of occlusion or highlights, we've proposed the *partial correlation* method to overcome the difficulty that encounters other methods. We've compared our method with the standard method and non standard method.

Why our technique is interest and important ? There are many good modeling techniques for discontinuities applied to the segmentation problem ([4]) but there are poor in modeling the discontinuities for multiple views. Our technique is a new attempt to solve this point. Intille [12] demonstrated the use of GCP (Ground control points) for building the disparity path along an epipolar line. For this purpose it is requiered that these points are as close as possible to the occlusion boundaries. For this reason, it's important to have a robust technique to match the points at the boundaries of objects, where *partial occlusion* occurs.

Experiments show that our method works better than other standard and non standard correlation methods when occlusion arises. However even if satisfactory results are obtained in non occluded areas, this method is outperformed by more standard correlations which run faster. Notice that for such detection partial correlation already provides some limits about occlusion position : the points which border the outliers region might be good candidates. Investigation on this approach is presently under work.

Acknowledgements

We would like to thank Pascal BRAND and Paolo REMAGNINO for their helpful input.

References

[1] P. Aschwanden and W. Guggenbühl. Experimental results from a comparative study on correlation-type registration algorithms. In Förstner and Ruwiedel, editors, *Robust Computer Vision*, pages 268–282. Wichmann, 1992.

[2] N. Ayache. *Stereovision and sensor fusion.* MIT-Press, 1990.

[3] D.H. Ballard and C.M. Brown. *Computer Vision.* Prentice Hall, 1982.

[4] A. Blake and A. Zisserman. *Visual Reconstruction.* The MIT Press, Cambridge, Massachusetts, 1987.

[5] P. Fua. A parallel stereo algorithm that produces dense depth maps and preserves image features. *Machine Vision and Applications*, 1990.

[6] D. Geiger, B. Ladendorf, and A. Yuille. Occlusions and binocular stereo. In G. Sandini, editor, *Proceedings of the 2nd European Conference on Computer Vision, Santa Margherita Ligure, Italy*, pages 425–433. Springer Verlag, 1992.

[7] F.R. Hampel, E.M. Ronchetti, P.J. Rousseeuw, and W.A. Stahel. *Robust Statistics : the Approach Based on Influence Functions.* Wiley series in probability and mathematical. John Wiley and Sons, New York, 1986.

[8] R. Horaud and Th. Skordas. Stereo correspondence through feature grouping and maximal cliques. IEEE *Transactions on PAMI*, 11(11):1168–1180, 1989.

[9] B. Horn and B. Schunk. Determining Optical Flow. Ai-memo 572, MIT, 1980.

[10] B.K.P. Horn. *Robot Vision.* The MIT Press, 1986.

[11] P.J. Huber. *Robust Statistics*, volume IX of *Wiley.* John Wiley, New York, 1981.

[12] S.S. Intille and A.F. Bobick. Disparity-space images and large occlusion stereo. In *Proceedings of the 3rd European Conference on Computer Vision, Stockholm, Sweden*, pages 179–186. Springer-Verlag, 1994.

[13] E. L. Lehman. *Nonparametrics: statistical methods based on ranks.* Holden-Day, 1975.

[14] D. Marr. *Vision*. W.H. Freeman and Company, San Francisco, California, USA, 1982.

[15] P. Meer, D. Mintz, A. Rosenfeld, and D.Y. Kim. Robust regression methods for computer vision: a review. *International Journal of Computer Vision*, 6(1):59–70, 1991.

[16] R. Mohr, P. Brand, and P. Remagnino. Correlation techniques in adaptive template matching with uncalibrated cameras. In *Vision Geometry III, SPIE's international symposium on photonic sensors & control for commercial applications*, volume 2356, pages 252–253, October 1994.

[17] H.K. Nishihara. PRISM, a pratical real-time imaging stereo matcher. Technical Report Technical Report A.I. Memo 780, Massachusetts Institute of Technology, 1984.

[18] J.M. Odobez and P. Bouthemy. Estimation robuste multi-échelle de modèles paramétrés de mouvement sur des scènes complexes. *International Journal of Computer Vision*, 11:419–430, 1994.

[19] P.J. Rousseeuw and A.M. Leroy. *Robust regression and outlier detection*, volume XIV of *Wiley*. J.Wiley and Sons, New York, 1987.

[20] B. G. Schunck. Robust computational vision. In *Proceedings of the International Workshop on Robust Computer Vision, Seattle, Washington, USA*, pages 1–19, October 1990.

[21] P. Seitz. Using local orientational information as image primitive for robust object recognition. In *SPIE proceedings*, pages 1630–1639, 1989.

[22] J. Weng. Image matching using the windowed Fourier phase. *International Journal of Computer Vision*, 11(3):211–236, April 1993.

[23] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondance. In *Proceedings of the 3rd European Conference on Computer Vision, Stockholm, Sweden*, pages 151–158. Springer-Verlag, May 1994.

[24] J. Zhao. *Extraction d'information tri-dimensionnelle par stéréovision*. Thèse de doctorat, Université Paul Sabatier, Toulouse, July 1989.

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399