# ACCURATE AND ROBUST IMAGE ALIGNMENT FOR ROAD PROFILE RECONSTRUCTION

*Jean-Philippe Tarel*[*] *Sio-Song Ieng* *Pierre Charbonnier*

LCPC
58 Bd Lefebvre
75732 Paris Cedex 15, France

LRPC d'Angers
23, Avenue de l'Amiral Chauvin, BP 69
F-49136 Les Ponts de Cé, France

LRPC de Strasbourg
11, Rue Jean Mentelin, BP 9
F-67200 Strasbourg, France

## ABSTRACT

In this paper we propose a novel approach of the two-image alignment problem based on a functional representation of images. This allows us to derive a one-to-several correspondence, multi-scale algorithm. At the same time, it also formalizes the problem as a robust estimation problem between possible matches. We then derive an accurate, robust and faster version for the alignment of edge images. The proposed algorithm is developed and tested in the context of *off-line* longitudinal road profile reconstruction from stereo images.

***Index Terms—*** Image Alignment, Registration, Robust estimation, Multi-scale, Stereovision, 3D Reconstruction, Road reconstruction.

## 1. INTRODUCTION

In the field of automotive applications, accurate and robust 3D road reconstruction is a key point for localizing a vehicle with respect to other vehicles and obstacles. In particular, it is required for long-range detection, where the usual planar-road assumption is not valid. It is also of main importance for building more accurate road maps.

The problem of accurate road reconstruction for obstacles detection was tackled as in [1] using an intermediate local reconstruction of images edges in the $(u, v, disparity)$ space also used in [2]. We propose here a new, direct approach for 3D road reconstruction from stereo pairs considered as an image registration problem under an unknown, parametric road model. The proposed approach is more robust and more accurate since it globally accounts for all the information at hand, *i.e.* the left and right edge images. Moreover, the problem being set in a robust estimation framework, the approximate covariance matrix of the estimate can be obtained and used as an inverse confidence matrix, which is of main importance in practical applications [3].

The proposed alignment algorithm is a particular case of non-rigid registration using a parametric model selected to fit road profiles as well as possible. There exists many image alignment techniques, see [4] for a survey. Unlike many existing algorithms, the one we propose here does not rely on any Taylor expansion but implements an exact optimization. This leads to improved convergence properties. Moreover, our algorithm is based on one-to-several rather than one-to-one correspondences. This strategy outperforms the well-known Iterative Closest Point (ICP) algorithm , as experimentally demonstrated in [5]. For solving the minimization, a multi-resolution strategy is used to converge to an interesting local minimum, in the spirit of [5, 6]. Finally, the proposed algorithm can be seen as an easier-to-derive and more flexible generalization of the approach of [5].

This paper is structured as follows: we first present the proposed parametric road model (Sec. 2), and in Sec. 3, we derive an algorithm for estimating the model parameters from a stereo pair. In Sec. 4, the road model is extended to take into account the roll angle of the vehicle. Finally in Sec. 5, we present experiments that assess the accuracy and robustness of the proposed alignment algorithm.

## 2. PARAMETRIC ROAD MODEL

We assume a rectified stereo-vision geometry. The image coordinates of a 3D point of the scene, with coordinates $(x_v, y_v, z_v)$ in the reference system of the vehicle, are:

$$\left.\begin{array}{rcl} \frac{u-u_0}{\alpha_u} & = & \frac{x_v+0.5b}{\cos\theta y_v-\sin\theta z_v} \\ \frac{u'-u'_0}{\alpha'_u} & = & \frac{x_v-0.5b}{\cos\theta y_v-\sin\theta z_v} \\ \frac{v-v_0}{\alpha_v} = \frac{v'-v'_0}{\alpha'_v} & = & \frac{-sin\theta y_v-cos\theta z_v+\frac{h}{cos\theta}}{\cos\theta y_v-\sin\theta z_v} \end{array}\right\} \quad (1)$$

where $(u, v)$ are coordinates within the left image, $(u', v')$ within the right image, $\alpha_u$, $\alpha'_u$, $\alpha_v$, $\alpha'_v$, $u_0$, $u'_0$, $v_0$, $v'_0$ are the intrinsic parameters of the two cameras, $b$ the stereo-vision base, $h$ the height of the cameras, and $\theta$ their pitch angle, see Fig. 1.

In practice, the left and right views mostly contain the image of the road. The shape of the road in front of the vehicle defines a mapping between the two images. The proposed approach consists of aligning the left image with the right one

[*]Thanks to the French Department of Transportation for funding, within the SARI-PREDIT project (http://www.sari.prd.fr/HomeEN.html).

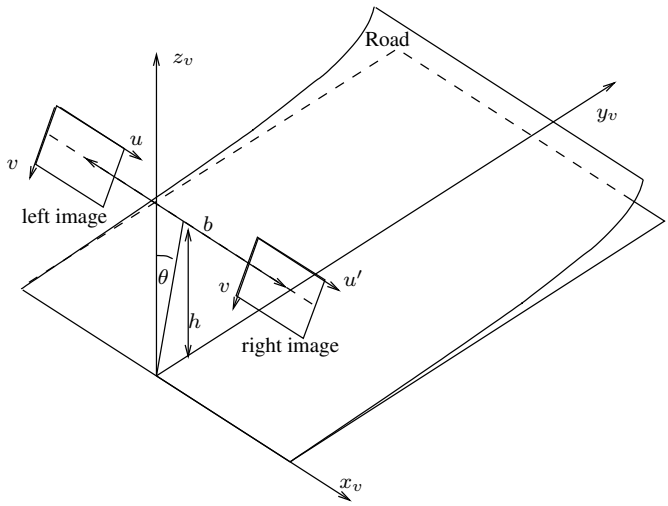



**Fig. 1**. Stereo geometry and road model.

consistently with an unknown, parametric road model. For a planar road model, *i.e.* $z_v = 0$, it is easy to show from (1), that the obtained mapping can be parametrized as $u' = u + a_0 + a_1 v$ and $v' = v$, where $a_0$ and $a_1$ are related to the intrinsic cameras parameters as well as to the position and orientation of the road plane. The observed road is not always planar, however. To tackle longitudinal variations of the road, we extend the previous mapping as:

$$\left.\begin{array}{rcl} u' & = & u + a_0 + a_1 v + \cdots + a_n v^n = u + A^t V(v) \\ v' & = & v \end{array}\right\} \quad (2)$$

where $V(v) = (1, v, v^2, \cdots, v^n)^t$ is the vector of monomials and $A = (a_0, \cdots, a_n)$ is the vector of (unknown) parameters related to the longitudinal road profile. We experimented several possible models of the profile, and it appears that a polynomial curve of degree $n$ is convenient enough, even though other functions might be considered as well. The advantage of the proposed parametric road model is its linearity with respect to the parameters, which allows us to derive without any approximation the algorithm described in the next section.

## 3. IMAGE ALIGNMENT FOR ROAD PROFILE RECONSTRUCTION

The two-image alignment problem is classically set as the minimization w.r.t. the mapping parameters of the squared errors between the left image $I_l$ after alignment, and the right image $I_r$ which is used as a reference. The first key problem within this approach is that images are discrete. It is thus necessary to interpolate the image $I_l$. Thanks to rectification, the interpolation must only be performed along the $u$ axis. The second key problem is that the optimization criterion may have flat areas, which prevents gradient descent or other usual numerical algorithms from converging. To bypass these problems, we propose to represent the digital right image $I_r(i, j)$ as the continuous function along $u$:

$$f_r(u, j) = \sum_i I_r(i, j) e^{-(\frac{i-u}{s})^2}$$

where $s$ is a scale parameter. After mapping, we similarly interpolate the digital left image $I_l(k, j)$:

$$f_l(u, j) = \sum_k I_l(k, l) e^{-(\frac{k + A^t V(j) - u}{s})^2}$$

Now, the squared error between the two images is given by:

$$e(A) = \frac{1}{2} \sum_j \int_u (f_l(u, j) - f_r(u, j))^2 du \quad (3)$$

By expanding the square under the integral, the error $e(A)$ can be simplified by forgetting the additive and multiplicative constants and written as minus the correlation of $f_l$ with $f_r$. After substitution of the expressions of $f_l$ and $f_r$, expansion and integration, a simpler expression of $e(A)$ is derived:

$$e(A) = -\sum_{i,j,k} I_l(i, j) I_r(k, j) e^{-\frac{1}{2s^2}(k + A^t V(j) - i)^2} \quad (4)$$

To minimize (4), with respect to $A$, we take advantage of the half-quadratic approach [7, 8] to derive an iterative algorithm without any Taylor expansion. Using [3] with $\phi(t) = -e^{-t}$, we check that $\phi$ fulfills the three required hypotheses: $\phi$ is defined and continuous on $[0, +\infty[$ as well as its first and second derivatives, $\phi$ is increasing and concave. Therefore the classical Kuhn and Tucker's theorem can be applied after introducing auxiliary variables, and the alternate minimization of the dual function leads to following iterative reweighted least-squares algorithm:

1. Initialize $A_0$, and set $p = 1$,
2. For all indexes $(i, j, k)$, compute the auxiliary variable:

   $$w_{i,j,k} = \frac{1}{2s^2}(k + A_{p-1}^t V(j) - i)^2$$

   and the Lagrange coefficient:

   $$\lambda_{i,j,k} = I_l(i, j) I_r(k, j) \phi'(w_{i,j,k})$$

3. Solve the linear system with respect to $A_p$:

   $$\sum_{i,j,k} \lambda_{i,j,k} V(j) V(j)^t A_p = \sum_{i,j,k} \lambda_{i,j,k} (i - k) V(j)$$

4. If $\|A_p - A_{p-1}\| > \epsilon$, increment $p$, and go to 2, else $A = A_p$.

More details on how the algorithm is formally derived can be found in [3]. Also, by applying results from [3], the convergence towards a local minimum is ensured. Moreover, by varying parameter $s$, a multi-scale strategy can readily be implemented to achieve convergence to a more interesting local minimum. This approach, called Graduated Non Convexity (GNC), consists of, at first, enforcing local convexity on the data using a large value of $s$. Then, a sequence of solutions with decreasing $s$, is computed in continuation, as in [6, 5].

It is important to notice that, in the obtained algorithm, the matching is one-to-several rather than one-to-one. As experimented in [5], one-to-several correspondence provides better convergence towards an interesting local minimum, and thus outperforms ICP or other one-to-one correspondence algorithms. This improved convergence property is detrimental in terms of computational burden. To avoid penalizing the proposed algorithm, we perform a matching decimation in the spirit of [5]: correspondences such as $k+A^tV(j)-i$ is larger than $3s$ are discarded. This decimation is performed only one time at each scale without a significant lost of accuracy.

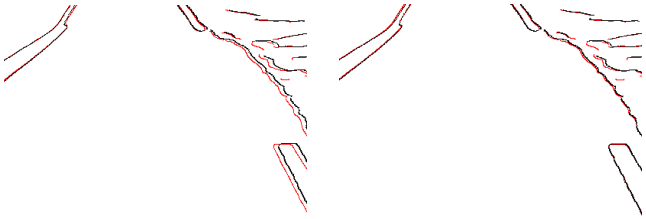
There are several ways to reduce the complexity of the proposed algorithm. For example, noticing that regions of constant intensity do not help the alignment process, it is faster to align edge maps rather than the original intensity images: we only have to work on possible edge matches $(i,k)$ for each line $j$. This leads to drastic reduction in complexity without lost of accuracy, as shown in Sec. 5. With edge maps, we can thus substitute the following expression to (4):

$$e(A) = \sum_{(i,k),j} d((i,k),j)\phi(\frac{1}{2s^2}(k + A^tV(j) - i)^2) \quad (5)$$

where $d((i,k),j)$ is an extra factor that is introduced to take into account the local similarity in terms of gray or color gradients between the two matched edge pixels $(i,j)$ and $(k,j)$. The proposed algorithm can thus be rewritten in a much more efficient way: at first the list of all possible matches is build for a scale $s$, and second the weights and $A_p$ are refined iteratively with $\lambda_{i,j,k} = d((i,k),j)\phi'(w_{i,j,k})$. The main difference compared with [5] is that the proposed algorithm is much more general in the choice of the weight $\lambda_{i,j,k}$. It is not reduced to the use of Gaussian weights and thus other robust functions may be used, which opens the door to using the proposed alignment algorithm in many applications with different kinds of perturbations.

## 4. ROLL ANGLE

Using the road model introduced in Sec. 2, we observed on several stereo pairs, situations in which one lane of the road is correctly aligned, while the reverse one is aligned with a few pixels bias, as shown on the rightmost part of Fig. 2(a). This is due to the fact that the roll angle of the vehicle is set to zero in the model. We now propose an extended road model featuring an extra parameter that accounts for the roll angle.



(a) (b)

**Fig. 2**. Alignment without (a) and with (b) roll angle estimation. Left edges are in red and the aligned right edges in black on this image crop.

When the road is planar, the roll slope $\lambda$ can be introduced in (1) by setting $z_v = \lambda x_v$. After this substitution, we simply derive the expressions of $x_v$ and $y_v$ as functions of $u$,$u'$, and $v$ from the first two equations of (1). Substituting these expressions in the last equation of (1) and simplifying, we obtain:

$$\left.\begin{array}{rcc} u' & = & u + a_0 + a_1 v + au \\ v' & = & v \end{array}\right\} \quad (6)$$

where $a$ accounts for the roll angle. Height variations of the road can now be introduced, in a similar way as in Sec. 2. The introduction of the roll angle still yields a mapping which is linear with respect to the unknown parameters. The algorithm proposed in Sec. 3 thus readily applies by changing $A^t$ into $(A^t, a)$ and $V^t(j)$ into $(V^t(j), u)$. The result obtained using the model including the roll angle on the same image is shown in Fig. 2(b). Notice how the alignment is improved for the edges on the right side of Fig. 2(b).

## 5. EXPERIMENTS

The proposed algorithm can be interpreted as a robust estimation algorithm on the matched pair points. The function that plays the role of the M-estimator is $\phi$ which can be chosen to achieve the maximum breakdown point of $50\%$ outliers, see [9]. Fig. 3 illustrates how robust the algorithm is in presence of many outliers due to vehicles, houses or trees. In the first column, a planar model is used, while in the second column it is a $6^{th}$ degree polynomial model. The advantage of the non-planar model is easily noticeable in Fig. 3.

To assess the accuracy of the alignment algorithm, we run experiments on synthetic images for different values of the transformation parameters, see Fig. 4. Note that the obtained accuracy is better using edge images than using gray level images. The correlation using edges is $96\%$ with a bias of about half a pixel. Typically, multi-scale alignment is performed in less than one hundred iterations and in a few seconds.

Finally, we address the problem of road profile reconstruction. We process a sequence of 330 stereo pairs taken at evenly spaced intervals along a 1650 meter long road. The true profile, obtained from the map is shown in blue on Fig. 5.

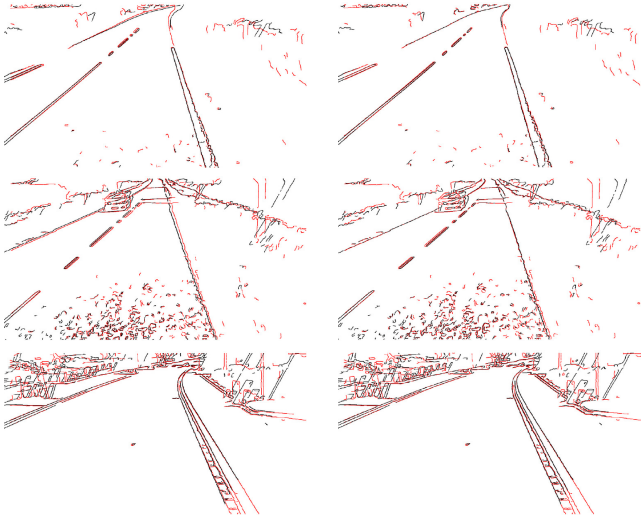
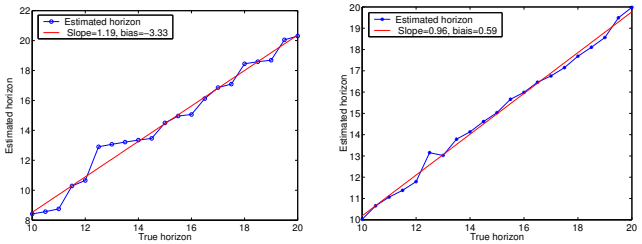


**Fig. 3**. Stereo registration using a planar model (left column) and with a $6^{th}$ degree polynomial (right column). The roll angle correction is incorporated. Black edges are the left edges mapped in the right image coordinates. Right edges are shown in red.



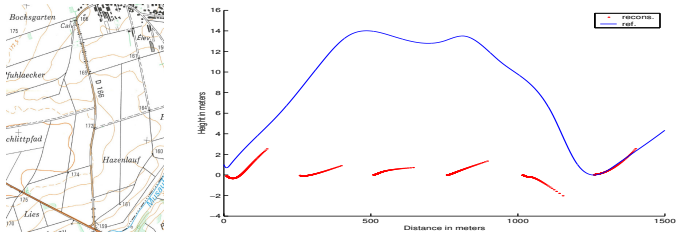

(a) (b)

**Fig. 4**. Estimated versus true horizon using gray level images (a) or edge images (b).

In red are shown 6 examples of local road profiles, each one resulting from the application of the algorithm on a particular stereo pair. Note that there is no integration of the profiles along the road path: they are treated individually. Their initial slope thus depends on the local orientation of the vehicle with respect to the road, which is unknown. Note, however that the relative shape of the local profiles is correct in all cases.

## 6. CONCLUSION

In this paper we derived a new algorithm for two-image alignment where one-to-several correspondences naturally arise and which is multi-scale. These two properties are of a main importance to achieve robustness of the alignment algorithm. The original formulation being relatively time-consuming, we propose a faster version by considering edge images. The alignment procedure is thus interpreted as a robust estimation over all possible edge pixels matches. Experimental results




(a) (b)

**Fig. 5**. (a) Map of the road (© IGN), (b) its longitudinal profile (in blue) and examples of estimated (relative) local road profiles (in red), see text.

show that this strategy is beneficial in terms of accuracy. We used with success this alignment algorithm for off-line longitudinal road profiles reconstruction from stereo images. Finally, the proposed approach is flexible enough to be easily extended to many other alignment problems.

## 7. REFERENCES


[1] R. Labayrade, D. Aubert, and J.-P. Tarel, "Real time obstacle detection in stereo vision on non-flat road geometry through v-disparity representation," in *IEEE Intelligent Vehicle Symposium (IV'2002)*, Versailles, France, 2002.

[2] J.-P. Tarel, "Global 3D planar reconstruction with uncalibrated cameras and rectified stereo geometry," *International Jounal Machine Graphics and Vision*, vol. 6, no. 4, pp. 393–418, 1997.

[3] J.-P. Tarel, S.-S. Ieng, and P. Charbonnier, "Using robust estimation algorithms for tracking explicit curves," in *European Conference on Computer Vision (ECCV'02)*, Copenhagen, Danmark, 2002, vol. 1, pp. 492–507.

[4] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, October 2003.

[5] S. Granger and X. Pennec, "Multi-scale EM-ICP: A fast and robust approach for surface registration," in *European Conference on Computer Vision (ECCV 2002)*, Copenhagen, Denmark, 2002, vol. 2353 of *LNCS*, pp. 418–432, Springer.

[6] J. M. Odobez and P. Bouthemy, "Robust multiresolution estimation of parametric motion models," *Journal of Visual Communication and Image Representation*, vol. 6, no. 4, pp. 348–365, 1995.

[7] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 3, pp. 367–383, 1992.

[8] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Transactions on Image Processing*, vol. 6, no. 2, pp. 298–311, 1997.

[9] S.-S. Ieng, J.-P. Tarel, and P. Charbonnier, "Modeling non-gaussian noise for robust image analysis," in *Proceedings of International Conference on Computer Vision Theory and Applications (VISAPP'07)*, Barcelona, Spain, 2007.