

# NON-LINEAR OPTIMIZATION FOR ROBUST ESTIMATION OF VANISHING POINTS

*Marcos Nieto and Luis Salgado*

Grupo de Tratamiento de Imágenes - Universidad Politécnica de Madrid - Madrid (Spain)  
{mnd, lsa}@gti.ssr.upm.es

## ABSTRACT

A new method for robust estimation of vanishing points is introduced in this paper. It is based on the MSAC (M-estimator Sample and Consensus) algorithm and on the definition of a new distance function between a vanishing point and a given orientation. Apart from the robustness, our method represents a flexible and efficient solution, since it allows to work with different type of image data, and its iterative nature makes better use of the available information to obtain more accurate estimates. The key issue of the work is the proposed distance function, that makes the error to be independent from the position of an hypothesized vanishing point, which allows to work with points at the infinity. Besides, the estimation process is guided by a non-linear optimization process that enhances the accuracy of the system.

The robustness of our proposal, compared with other methods in the literature is shown with a set of tests carried out for both synthetic data and real images. The results show that our approach obtain excellent levels of accuracy and that is definitely robust against the presence of large amounts of outliers, outperforming other state of the art approaches.

*Index Terms*— Vanishing point, MSAC, projective plane.

## 1. INTRODUCTION

Vanishing point estimation is a major topic in the computer vision research community since decades, and for which there exist numerous excellent works [1][2][3][4]. In fact, from a purely theoretical point of view, solutions to the problem have been already reached, since there are a number of methods that are able to determine the position of multiple vanishing points through the use of image information, without assuming any a priori information [3]. Nevertheless, from the practical point of view, there are many aspects to be solved. The trend now is to propose new contributions not only for improved accuracy, but for efficiency (how much information is required by the method?), flexibility (what type of information can it use?), automatism (can it be implemented for any computer vision task?) and speed (can it be really used for online systems?).

Recent works are typically based on optimization processes that face the problem from a Bayesian point of view [1]-[3]. In this framework the existence of one or several vanishing points is modeled as a random variable not directly observable, but from which noisy measurements are available. One of the most interesting aspects to be solved in this type of approaches is how to make the estimations robust against outliers. Typically, the complexity of the models is increased to consider that the image data set is formed by

both inliers and outliers. Nevertheless, the presence of outliers introduces a high degree of uncertainty in the estimation process due to their unpredictable nature. Fortunately, the Bayesian framework has been shown to adequately treat this uncertainty quite naturally so that robust approaches typically use methods that are Bayesian at heart, such that the EM algorithm [1][3][4] or RANSAC [2][5].

Nevertheless, these works have paid more attention to aspects like accuracy, completeness or adequacy of the noise models rather than to the final robustness of the estimation system against significant amount of outliers.

For this reason, in this work we propose a new vanishing point estimation strategy, that has been specially designed to be highly robust against large amounts of outliers. This is the typical case of images of real structured scenarios with additional noisy elements such as persons, vehicles, wires, trees, shadows, or any other element that is not oriented towards a significant vanishing point. Our approach is built on the well known MSAC extension of the RANSAC algorithm (M-estimator Sample and Consensus [6]). This re-descending estimator controls much more tightly the impact of the outliers in the computation of the cost of the consensus set delivering much clearer and better results in terms of accuracy and efficiency.

A significant advantage of our algorithm is its efficiency against more complex strategies, such as EM-based ones [3]. In the Bayesian framework, with ML or MAP approaches, the main drawback of EM-based strategies is that the complete data set must be considered when computing the M-step at each iteration. This is definitely much more costly than random sampling strategies like MSAC, in which limited subsets of data are chosen sequentially to generate hypotheses. Besides, it is possible to reach convergence in few iterations by using an appropriate sampling strategy as we propose in this work.

An error or distance function must be defined in the MSAC framework, such that the optimization processes search for the value of the parameters that minimize a global cost function. The proposed distance function models the error as the sine of the angular deviation between the expected orientation and the measured orientation of the data samples. This proposal is guaranteed by the fact that the orientation is the main source of information for computing vanishing points [3][7] whereas it is independent on the position of the vanishing point, allowing to handle those at infinity. Besides, our formulation includes weighting factors to the distance function in order to drive the estimation process considering more relevant the contribution of longer line segments or, in case of using pixel gradients as the data set, those showing higher gradient magnitude.

The proposed method reaches excellent results in terms of robustness, even in the presence of large amounts of outliers, and accuracy, due to the non-linear optimization process included into the MSAC structure. Besides, the method is flexible and efficient, due to its ability to adapt to the type of image data and the iterative nature of the algorithm.

---

This work has been supported by the Ministerio de Ciencia e Innovación of the Spanish Government under projects TEC2007-67764 (SmartVision) and TEC2006-26845-E (Highway), and by the European Commission 6th FP under project IST-2004-027195 (I-WAY).

## 2. MSAC STRATEGY

From here on we will assume that the image features are line segments, since most of the works we will compare with use them as source information [1][2][8].

The estimation of vanishing points is carried out with an iterative scheme, in a similar way to other works [2]. At each iteration, given the complete set of observations  $\mathcal{L} = \{\mathbf{l}_i\}_{i=1}^N$ , a single dominant vanishing point is hypothesized, such that each line segment,  $\mathbf{l}_i$ , is assumed to correspond to one of two classes, that are labeled as  $c_i = \{ON, OFF\}$ , indicating respectively that the line segment is oriented towards the vanishing point or not. The likelihood model, assuming i.i.d. samples is defined as  $p(\mathcal{L}|\mathbf{v}) = \prod_{i=1}^N p(\mathbf{l}_i|\mathbf{v})$ ; where  $p(\mathbf{l}_i|\mathbf{v}) = \sum_{c_i} p(c_i)p(\mathbf{l}_i|\mathbf{v}, c_i)$ , and the density  $p(\mathbf{l}_i|\mathbf{v}, c_i = ON)$  is modeled as a normal distribution:

$$p(\mathbf{l}_i|\mathbf{v}, c_i = ON) \propto \exp\left(-\frac{1}{2\sigma^2}d^2(\mathbf{l}_i, \mathbf{v})\right) \quad (1)$$

where  $d(\mathbf{l}_i, \mathbf{v})$  is the distance function proposed in this work, explained in Section 3. After the estimation, the inliers are removed from the data set and the presence of another dominant vanishing point is hypothesized.

Summarizing, the MSAC algorithm proceeds iteratively in two main steps: (i) a minimal sample subset of data is randomly selected in order to generate a hypothesis of the vanishing point. In this case, the minimal set is composed by two line segments whose best estimation is  $\mathbf{v}^* = \mathbf{l}_i \times \mathbf{l}_j$ ; and (ii) the consensus set,  $s(\mathbf{v}^*)$ , i.e. the subset of line segments that are coherent with the hypothesis is computed as  $s(\mathbf{v}^*) \triangleq \{\mathbf{l}_i \in \mathcal{L} : d^2(\mathbf{l}_i, \mathbf{v}^*) \leq \delta\}$ ; where  $\delta$  is a certain threshold governed by the statistics of the data (details are given in Section 3.2).

These two steps are repeated until the probability of finding a better consensus set is below a certain threshold. Here resides the difference between MSAC and RANSAC: while basic RANSAC ranks the consensus sets according to their cardinality, i.e. the number of elements of the set, the MSAC version ranks them according to a global cost function that gathers individual contributions of inliers and outliers. The total cost of an hypothesis of the MSAC iterative procedure is given by  $\mathcal{C}(\mathcal{L}, s(\mathbf{v}^*)) = \sum_{i=1}^N \rho(\mathbf{l}_i, \mathbf{v}^*)$ , where  $\rho(\mathbf{l}_i, \mathbf{v}^*)$  is defined as:

$$\rho(\mathbf{l}_i, \mathbf{v}^*) = \begin{cases} d^2(\mathbf{l}_i, \mathbf{v}^*) & , d^2(\mathbf{l}_i, \mathbf{v}^*) \leq \delta \\ \delta & , \text{otherwise} \end{cases} \quad (2)$$

As a final remark, the minimal sample subset selection can be carried out more efficiently by using any type of sampling strategy that is more peaked than a purely random process. In this work, we propose to link the probability of a line segment to be selected to its length, since longer line segments are more likely to be less noisy [2] and, therefore, more likely to be inliers.

## 3. DISTANCE FUNCTION

This section describes the proposed distance function  $d(\mathbf{l}_i, \mathbf{v})$ , including a brief overview of the advantages of its use compared with other proposed distances in the literature. Further subsections show the integration details of this function into the MSAC structure, mainly in terms of the ranking process.

### 3.1. Orientation-based error function

Many different error functions have been proposed in the literature to characterize the error between an image feature (typically a line

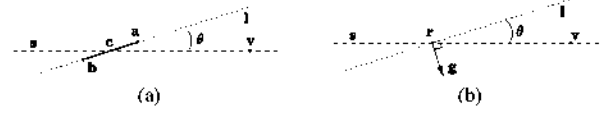


Fig. 1. Relative orientation between the orientation of the measured feature and the vanishing point ((a) line segments and (b) edge-points). When the vanishing point lies on  $\mathbf{l}$ , then  $d(\mathbf{v}, \mathbf{l}) = 0$ .

segment), and a vanishing point. The vanishing point to line error model, i.e. point-line distances in the image plane [5][9], are the simplest ones, but only can be applied when the vanishing point is close to the image. For instance, these distances are not valid for vanishing points at the infinity, as the further a vanishing point is from the line segment, the larger the error will grow. Though, if the calibration matrix is known, the coordinates of the image plane can be normalized to the unit sphere. Hence, the point-line distance becomes a relative difference between two 3D orientations [1].

As a specific solution for line segments, defined by a pair of end-points, Liebowitz [8] proposed to construct a cost function based on the sum of the point-line distance of each end-point with respect to a line that joins the vanishing point and the line segment. This approach also accounts for an optimization process in which this line is dependent on the position of the vanishing point, minimizing the global committed error.

Nevertheless, more general approaches [3][7] consider the orientation error as the workspace on which to perform optimization such that the error function can be applied indifferently with pixel gradients and line segments. For instance, Schindler [3] considers the angle between the normal orientation of a pixel gradient and the line joining the vanishing point and that pixel. Inspired by this approach, we propose to use the sine of the difference between the measured orientation of the line segment and the hypothesized vanishing direction, in homogeneous coordinates.

Let's consider the distance between a vanishing point,  $\mathbf{v}$  in homogeneous coordinates and the line  $\mathbf{l}$  as the absolute value of the sine of the angle between  $\mathbf{l}$  and a line,  $\mathbf{s}$ , joining  $\mathbf{v}$  and a reference point of the image feature (line segment or pixel gradient)

$$d(\mathbf{l}, \mathbf{v}) = \left| \frac{-l_2 s_1 + l_1 s_2}{\sqrt{l_1^2 + l_2^2} \sqrt{s_1^2 + s_2^2}} \right| \quad (3)$$

In the case of using line segments, the line  $\mathbf{s} = (s_1, s_2, s_3)^T$  is computed as  $\mathbf{s} = \mathbf{v} \times \frac{1}{2}(\mathbf{a} + \mathbf{b})$ , and  $\mathbf{l} = \mathbf{a} \times \mathbf{b}$ , where  $\mathbf{a}$  and  $\mathbf{b}$  are the end-points of the line segment. In the case of edge-points, which are defined by the duple of point and gradient vector  $\{\mathbf{r}, \mathbf{g}\}$ , the lines are defined as  $\mathbf{s} = \mathbf{v} \times \mathbf{r}$ , and  $\mathbf{l}^T \mathbf{g} = 0$  in homogeneous coordinates. These concepts are illustrated in Fig. 1.

The absolute value is considered because we are only interested in the relative deviation between the orientation of the lines, not in its sign, so that the angle between them is limited between 0 and  $\pi$ . The usage of the sine function is motivated by the fact that, for low values of the deviation angle, which is the case for inliers,  $\sin \theta \simeq \theta$ .

In order to take into account other sources of information, our approach can be extended by adding a scale factor to (3) that weights the importance of image features in the optimization process. The distance is modified as:

$$d'(\mathbf{l}_i, \mathbf{v}) = S \cdot d(\mathbf{l}_i, \mathbf{v}) \quad (4)$$

where  $S$  could be any type of factor that weights the importance of image features. For instance, in the case of using edge-points,

$S$  could be the inverse of the magnitude of their gradient, so that more intense edge-points are more important. For line segments, this could be related to their length or any quality measurement. In this work we propose to use the length of the line segments as scale factor. As it will be shown, the use of this information makes the system more robust against outliers.

### 3.2. Estimation step

Within the MSAC framework, the hypothesis and test steps (described in Section 2), can be thought of as an equivalent to the E-step of the EM algorithm, since they comprise the selection of a minimal sample subset, as well as the computation of the corresponding consensus set.

The threshold  $\delta$ , that divides line segments between inliers and outliers, given a vanishing point hypothesis, is selected as  $p(d^2(\mathbf{l}_i, \mathbf{v}) \leq \delta) = P_{\text{inlier}}$ . It can be computed as  $\delta^2 = \sigma^2 F_{\chi^2_n}^{-1}(P_{\text{inlier}})$ . The value of  $P_{\text{inlier}}$  is typically set to 95%, and  $\sigma$  is the expected standard deviation of the noise of the inliers. In the case of one-dimensional noise,  $n = 1$ , (as for  $d$ ), assuming that the maximum deviation angle is  $\theta_{\text{max}} = \pi/16$ , and that  $2.5\sigma = \sin(\pi/16)$  swaths the 95% of the probability, then  $\sigma = 0.065$  and  $\delta = 0.01623$ .

### 3.3. Maximization step

Following the analogy of the EM algorithm, we can define the M-step of the MSAC strategy as the re-estimation of the vanishing point considering the contribution of all the line segments of the consensus set. Provided that the conditional distribution (1) is normal, the maximum likelihood estimation is solved by minimizing the sum of squared costs  $\mathcal{C} = \sum_{i=1}^N d^2(\mathbf{l}_i, \mathbf{v})$ . However, this is a difficult problem that can not be solved analytically as the expression to maximize is highly non-linear. Therefore approximate numerical methods for non-linear optimization are required. In this work we propose to use the Levenberg-Marquardt algorithm.

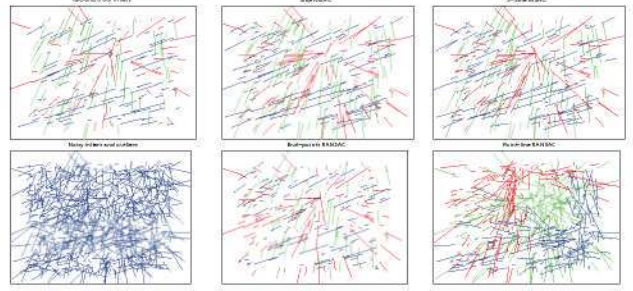
## 4. TESTS AND RESULTS

In this section we compare the accuracy and robustness of our approaches, that we will call Sine MSAC, and its extension using the scale factor, S-Sine MSAC, with two different methods of the literature: the RANSAC using the Liebowitz error function [2], and the RANSAC algorithm using the normalized homogeneous point-line distance (which is similar to the work by [1], although they use the EM algorithm). Both synthetic and real examples are used in order to provide a complete set of results.

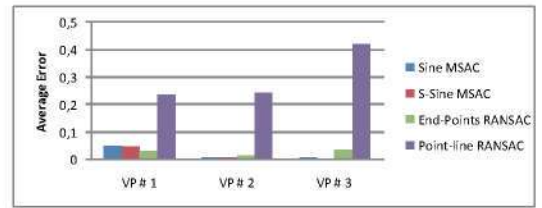
### 4.1. Synthetic data

Fig. 2 shows three sets of ground truth line segments concurring into three vanishing points. The test consists on adding noise to these line segments and an additional set of outliers, generated as random line segments inside the image boundaries. As shown, the performance of Sine MSAC, S-Sine MSAC and end-points RANSAC is very similar and very close to the ground truth for all vanishing points. The point-line RANSAC method is finally not robust against this amount of outliers, and its estimations are far from correct.

This experiment is repeated varying the number of outliers, ranging from 0 to 1000, and the results are shown in Fig. 3 as the average error committed by each method. In this case, the results are shown separately for each vanishing point, and the average error is



**Fig. 2.** Example of behavior of the robust strategies in the presence of noisy inliers and a large set of outliers: The number of inliers is  $\{50, 100, 100\}$  respectively for each vanishing point, and the number of outliers is 600.



**Fig. 3.** Results of the proposed robust methods for different number of outliers. In (a) the three ground truth line segments and vanishing point are shown separately; (b) shows the corresponding error values according to the number of outliers; and (c) shows the average error.

shown for each method, computed as the sine of the angle between the ground truth vanishing point and the estimated vanishing point.

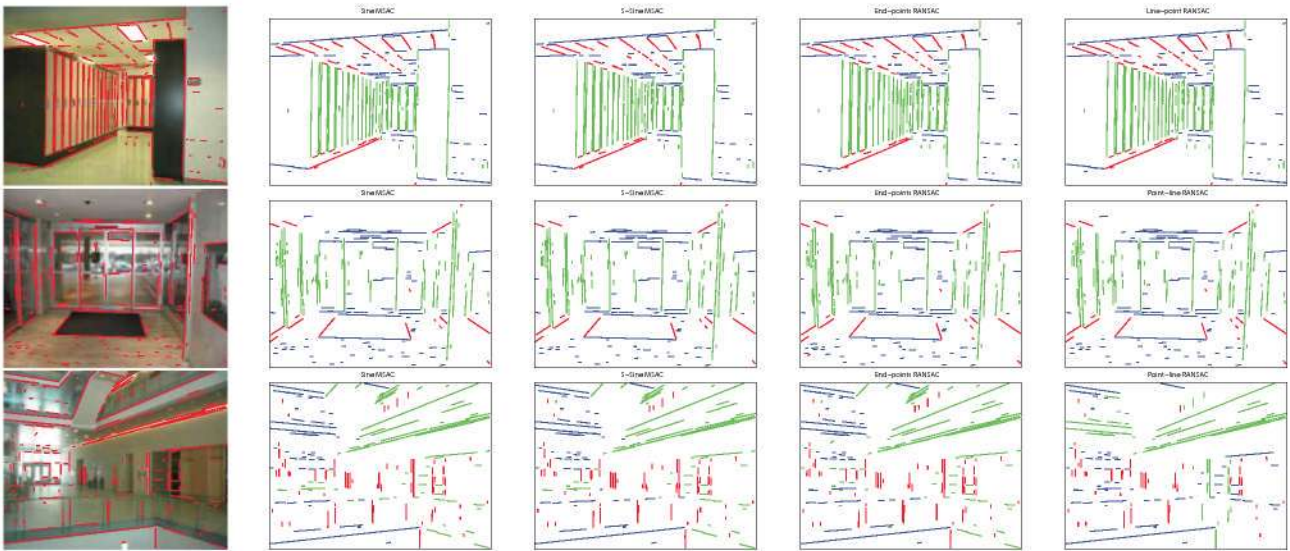
From this figure, the immediate conclusion is that the point-line method fails even for a small number of outliers, while the MSAC-based methods offer good results up to certain proportion of outliers. Particularly, the Sine MSAC and S-Sine MSAC methods obtain errors below 0.05, which is very accurate (less than 3 degrees), for up to 400 outliers (for 50 noisy inliers) in the worst case, and up to 1000 for the other two vanishing points. The reason is that the first case corresponds to a vanishing point inside the boundaries of the image, which is much more affected by the presence of random outliers in the image. As summary, the MSAC-based methods perform very well for all the vanishing points, being the S-Sine MSAC method the more accurate in average.

### 4.2. Real data

To test the performance of the proposed methods with real images we have chosen an available image database, called the York Urban Data Base [7], composed by 102 images of structured environments (corridor, facades, parkings, etc.) for which the calibration information is available as well as the ground truth of the three dominant vanishing points.

We measure the errors in angles, because we have the calibration of the camera, that makes that image plane positions of vanishing points become 3D vanishing directions. For clarity in the results, we consider that a method has correctly detected a vanishing point when the error angle is below 10 degrees with respect the ground truth.

For the extraction of the line segments we have designed a fast method that is based on a random sampling strategy that sequentially select pixels in the image according to their gradient magnitude. For



**Fig. 4.** Examples of the estimated vanishing points for different test images and different methods. At the left column the images are shown with the superimposed detected line segments. The color code of the line segments represents to which vanishing point they are oriented.

each selected pixel a non-maximal suppression and a growing algorithm based on orientation comparisons between neighbor pixels are applied in order to determine the line segment.

Attending to the results, it is noteworthy that the Sine MSAC, and specially the S-Sine MSAC obtain the best results in average. For instance, the rate of correct detections when three vanishing points are present are 80.7%, and 84.6%, respectively. The End-points RANSAC obtains 79.5%. The point-line RANSAC obtain slightly worsen results, with a rate of 69.2%.

Considering only the correct detections, the average error is very low for the four methods, without significant differences, that in fact may be due to inaccuracies in the ground truth vanishing points, or the effect of the radial distortion.

Fig. 4 shows the results of the different methods for several images of the used database. The first two images (from top to bottom) are examples of the excellent performance of the proposed methods when the image contain enough information for all vanishing points, which is the typical case of man-made environments. The last example show what could happen when there are other elements in the image, such as many line segments oriented towards different orientations, that may cause confusion to the estimation methods, as is the case for the point-line RANSAC, that fails to obtain the green vanishing point.

## 5. CONCLUSIONS

In this paper we have introduced a new method for robust estimation of vanishing points, that is based on the MSAC algorithm, and on the definition of a novel distance function between a vanishing point and an orientation. The approach is flexible in the sense that can be used with line segments and pixel gradients without difference, efficient taking into account the iterative nature of the MSAC algorithm, and accurate due to the used non-linear optimization process.

Excellent results have been obtained mainly in terms of robustness and accuracy, that have been shown with different tests using synthetic data and real images from a known image database. The proposed method, and an extension that makes use of additional a

priori information have been compared with other methods in the literature, outperforming their results and showing that our proposal is very robust even against large amounts of outliers.

## 6. REFERENCES

- [1] J. Kosecka and W. Zhang, "Video compass," in *European Conference on Computer Vision, LNCS 2350*, Springer Verlag, 2003, pp. 476–491.
- [2] R. Pflugfelder, *Self-calibrating cameras in video surveillance*, Ph.D. thesis, Graz University of Technology, May 2008.
- [3] G. Schindler and F. Dellaert, "Atlanta world: An expectation maximization framework for simultaneous low-level edge grouping and camera calibration in complex man-made environments," in *Proc. Conf. on Computer Vision and Pattern Recognition*, 2004, pp. 203–209.
- [4] M. E. Antone and S. Teller, "Automatic recovery of relative camera rotations for urban scenes," in *Proc. Conf. Computer Vision and Pattern Recognition*, 2000, vol. 2, pp. 282–289.
- [5] H.-H. Trinh and K.-H. Jo, "Image-based structural analysis of building using line segments and their geometrical vanishing points," in *SICE-ICASE International Joint Conference 2006*, Oct. 2006, pp. 566–571.
- [6] P.H.S. Torr and A. Zisserman, "Mlesac: A new robust estimator with application to estimating image geometry," *Journal of Computer Vision and Image Understanding*, vol. 78, no. 1, pp. 138–156, 2000.
- [7] Denis P., Elder J.H., and F. J. Estrada, "Efficient edge-based methods for estimating manhattan frames in urban imagery," in *Proc. ECCV08*, 2008, vol. 2, pp. 197–210.
- [8] D. Liebowitz, *Camera calibration and reconstruction of geometry*, Ph.D. thesis, University of Oxford, June 2001.
- [9] A. Minagawa, N. Tagawa, T. Moriya, and T. Gotoh, "Vanishing point and vanishing line estimation with line clustering," *IEICE Trans. Inf. & Syst.*, vol. E83-D, no. 7, July 2000.