Robust Homography Estimation from Planar Contours Based on Convexity

Alberto Ruiz[1], Pedro E. López de Teruel[2], and Lorenzo Fernández[2]

[1] Dept. Informática y Sistemas, University of Murcia
[2] Dept. Tecnología e Ingeniería de Computadores, University of Murcia
aruiz@um.es, {pedroe, lfmaimo}@ditec.um.es

Abstract. We propose a homography estimation method from the contours of planar regions. Standard projective invariants such as cross ratios or canonical frames based on hot points obtained from local differential properties are extremely unstable in real images suffering from pixelization, thresholding artifacts, and other noise sources. We explore alternative constructions based on global convexity properties of the contour such as discrete tangents and concavities. We show that a projective frame can be robustly extracted from arbitrary shapes with at least one appreciable concavity. Algorithmic complexity and stability are theoretically discussed and experimentally evaluated in a number of real applications including projective shape matching, alignment and pose estimation. We conclude that the procedure is computationally efficient and notably robust given the ill-conditioned nature of the problem.

1 Introduction

The homography relating two perspective views of a plane is a fundamental geometric entity in many computer vision applications. Instead of conventional estimation methods based on explicit point or line correspondences, we are interested in robust and efficient homography estimation from the *contours* of two views of a given planar region with arbitrary shape. Using this transformation we can solve several related problems including shape recognition and matching, object alignment, spatial pose location (given additional information about the camera parameters), robot guidance from conventional signs (e.g. arrows), image rectification and camera calibration.

For instance, Figs. 1.a-b show two views of a well-known geographical feature. Using the homography relating the two views we could verify that the aerial image effectively corresponds to the lake in the map, the cities in the map can be located on the image, and we can even compute the 3D position and orientation of the camera in the reference frame induced by the map.

These natural shapes lack distinguished points or lines; at a given resolution they can be considered just as irregular silhouettes in which small details are neither reliable nor relevant. Furthermore, contours extracted from real images suffer from pixelization, thresholding artifacts, and other unavoidable noise sources, specially in low resolution views with large slant (Fig. 1.c).

A. Leonardis, H. Bischof, and A. Pinz (Eds.): ECCV 2006, Part I, LNCS 3951, pp. 107–120, 2006.
© Springer-Verlag Berlin Heidelberg 2006

(a) (b) (c)

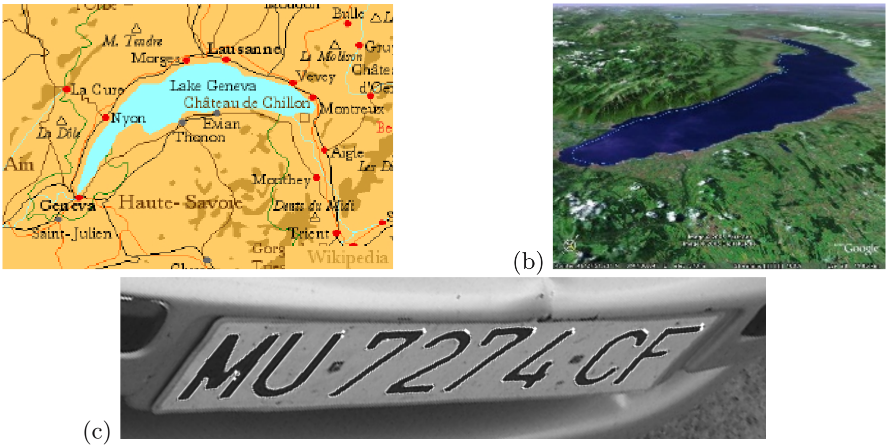


Fig. 1. Real world shapes. (a-b) Two views of Lake Geneva. (c) Noisy contours of traffic plate symbols extracted from a video sequence.

In noisy contours the differential properties of curves (required for computation of lines, inflection points, cusps, and other local projective invariants) are destroyed. Cross-ratio constructions are also very sensitive to noise and must be used with caution. Contour smoothing and noise filtering do not completely solve this problem: noise is inhomogeneously transmitted in different regions of the contour due to the nonlinear effects of perspective imaging. Analytical models (e.g. polygonal approximations, implicit polynomials, *snakes*, etc.) may even destroy valuable features for contour alignment. Certain modeling techniques may be adequate for specific shapes (e.g. straight line approximations for essentially polygonal contours, etc.), but contour recognition in general conditions is precisely one of our main goals. In consequence, in this paper all contours will be represented and manipulated in its "raw" form as closed and possibly irregularly spaced polylines without self-intersections.

Some of the first approaches to shape recognition under perspective imaging conditions were based on more or less *ad hoc* constructions [1]. Later, the application of projective geometry [2, 3, 4] to computer vision clarified enormously the field, but the emphasis was mainly in estimation of 3D structure from explicit point or line correspondences in multiple images.

Projective contour analysis under real world, noisy conditions has received comparably less attention. Most of the proposed solutions for curve matching are based on differential properties [5, 6, 7, 8, 9] or in specific contour models [10, 11], which cannot be directly used over low quality images The projective geometry of multiple views of curves has been studied in [12]. Invariant signatures based on rays have been proposed in [13] to retrieve shapes in a database of trademarks. Application of contour matching to visual servoing using snakes is described in [14], where weak perspective estimates, point redistribution, and projective correction steps are iterated until convergence. An approach based on image moments is reported in [15]. A curious and completely different idea is proposed in [16], where a linear program can be established on the homography entries, with constraints given by region bounds. This method admits partial occlusions but requires at least two contours to avoid trivial solutions. In addition to shape recognition, contour alignment has been used in other applications including camera calibration [17] and structure and motion recovery [18, 19].

Contour matching under similar or affine transformations (e.g. weak perspective) is a notably easier problem [20, 21, 22]. For instance, robust affine alignment can be based on shape covariance equalization and Fourier analysis. Unfortunately, this kind of approaches cannot be directly extended to full perspective images due to the essentially nonlinear laws of image formation. While small shapes can frequently be acceptably modeled by affine transformations, such kind of weak perspective approximation is only valid for shape recognition. Accurate alignment and pose estimation can only be achieved from true projective homographies containing information about both the focal length and the distance to the object.

Our goal is a simple, efficient, and robust method for homography estimation from arbitrary contours. In the rest of the paper we will discuss a number of geometric constructions, essentially based on convexity, which can be used to compute a projectively invariant reference frame.

2 Robust Projective Invariants

The homography relating two projective views of a plane is completely characterized by at least four corresponding points (or lines) [2]. However, two corresponding contours only impose (if differential or local properties are discarded) an ordering on the possible point correspondences. Distances between points along the contour may drastically expand or shrink in different views. We are interested in a projective reference frame that can be constructed using 'global' invariant geometric properties of the curve, avoiding local properties. The construction must be tolerant to a reasonable amount of noise in the curve locations.

A promising property is convexity. The convex hull of a figure is preserved under projective transformations if the whole shape is in front of the camera (otherwise objects are split across the horizon; we consider *quasi-affine* transformations [2, ch. 21], [23]). In this work we assume that the admissible contours are completely contained in the image, without occlusions. In such conditions, and in contrast with curvature-based invariants, the global convexity properties of a figure can only be destroyed by large amounts of noise. This kind of region

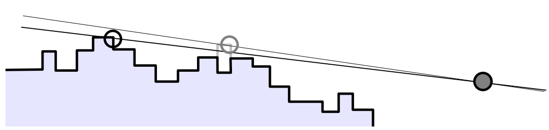


Fig. 2. The discrete tangent with respect to a external point (but not the point of contact) is reasonably robust against contour perturbations

convexity invariance seems to be a minimal and reasonable requirement. If 'large' concavities disappear contour matching becomes unsolvable in practice.

Closely related to convexity, tangency is also projectively preserved. While ordinary curve tangents, based on differential properties, are not robust, 'discrete' tangency with respect to external points or regions is a much more stable geometric construction (Fig. 2). Note that the specific point of contact *is not* a robust projective invariant (it may slide along the tangent line).

2.1 Polygon Tangency and Convex Hull Computation

The points of contact of the tangents to a polygon are contained in its convex hull, which can be efficiently computed using Melkman's algorithm for polylines with no self-intersections [24]. This method sequentially processes each of the polyline vertices. At each stage, the algorithm determines and stores on a double-ended queue those vertices that form the ordered hull for all polyline vertices considered so far. Each new vertex satisfies one of two conditions (Fig. 3): either (1) it is inside the currently constructed hull, and can be ignored; or (2) it is outside the current hull, and becomes a new hull vertex extending the old

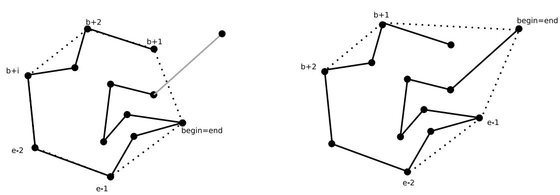


Fig. 3. Illustration of one step in Melkman's convex hull algorithm

hull. However, in case (2), vertices that are on the list for the old hull, may become interior to the new hull, and need to be discarded before adding the new vertex to the new list. Each vertex can be inserted on the deque at most twice (once at each end) and the elements on the deque can be removed at most once. Each of these events has constant time, providing a linear execution order.

2.2 Contour Pairs

To illustrate a simple example of convexity based invariants we will consider first the easiest situation. Given a *pair* of closed, disjoint coplanar contours, the four tangent lines to both contours is an eight d.o.f. projective invariant which completely determines the homography relating two views (Fig. 4).

This idea can be immediately applied to planar objects with at least two holes (e.g. the shape "B"), but obviously we are actually interested in the more general case of simple contours without holes. In principle, this method could be applied to figures with at least two clear concavities (which, together with the

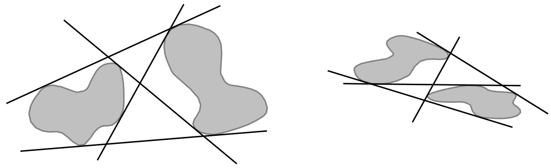


Fig. 4. Four invariant lines from a contour pair

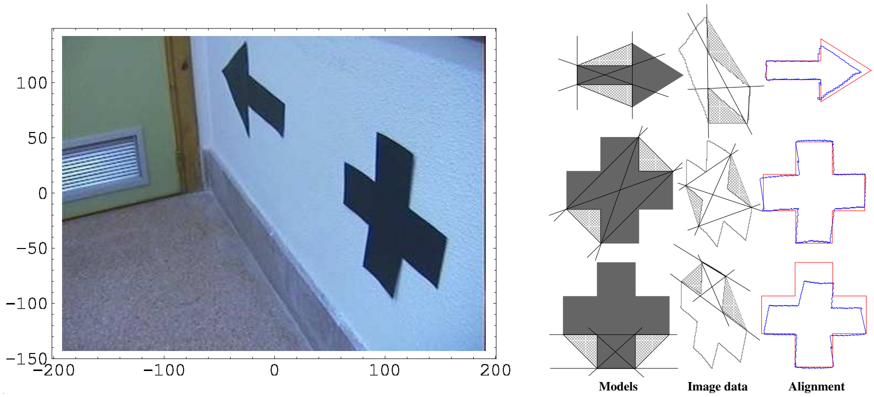


Fig. 5. Alignment using a pair of concavities

convex hull, are also projectively invariant). For example, Fig. 5 illustrates this idea for projective alignment of signs in a robot guidance application [25] [1].

Alignment is acceptable despite the bad quality of the signals, which are loosely glued to the wall. As shown in the last row, alignment quality strongly depends on the chosen pair of concavities: we must try all combinations and return the best match. Homography computation from the corresponding lines becomes ill conditioned if the contours in the pair are too close, or too separated, or their sizes are disparate.

In any case, this method is in general not robust since concavities are actually defined by *open* contours with extremes that may slide along the convex hull. The bitangent contact points induced by the concavities, which could in principle be used to define a projective reference frame, are also unstable. In the next section we propose a more robust and general alignment method.

3 Single Concavity

Under ideal conditions a smooth concavity defines at least four invariant points (Fig. 6.a) which specify a projective reference frame [5, 6] (the points supporting

[1] In this particular example polygonal models could directly provide candidate lines or vertices for matching. However, the proposed model is completely general and only the raw contours are required for alignment.

the concavity base line (bitangent) and the inflection points or the points of contact of tangents). However, these points are not stable in most real, noisy situations, and may even be not defined (Fig. 6.b).

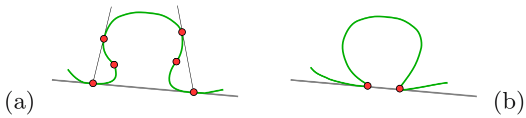


Fig. 6. Invariant points specified by a concavity

Consider instead projective frames defined essentially by discrete tangents. Disregarding local curvature, a convex shape can only reduce the 8 degrees of freedom of an arbitrary homography to 4, namely the angle/position of contact of four lines enclosing the shape (Fig. 7 (left)). Therefore, a *smooth* convex shape can robustly specify neither projective nor affine (6 d.o.f.) reference frames.

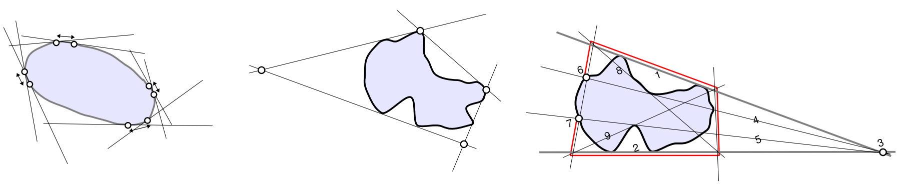
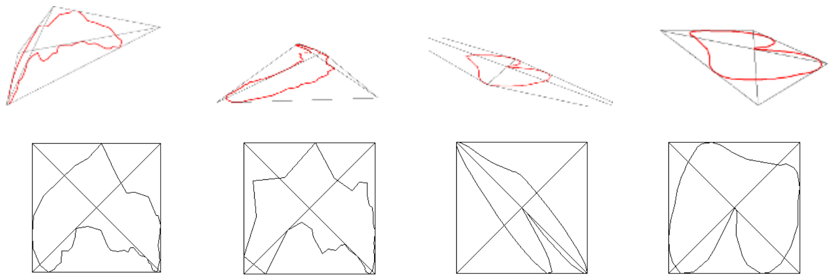


Fig. 7. Using only tangencies a convex shape can only fix 4 d.o.f. in a homography (left). Projective frame completely fixed using 4 (center) and 2 concavities (right).

We need some appreciable concavities (or straight line fragments) in the shape in order to constrain the remaining degrees of freedom of the projectivity with additional tangencies. The bitangent of a concavity is a robust invariant in the sense of Sect. 2 (clearly, its stability increases with the distance between the contact points). A convex shape with four or more concavities trivially defines one or more projective frames (Fig. 7 (center)). The bitangents are efficiently computed as a side effect of the convex hull algorithm. Interestingly, taking advantage of tangents to the concavities and intersections with the convex hull only two of them are actually required to define a projective frame (Fig. 7 (right)). Of course, many other alternative constructions can be conceived; practical considerations suggest that the most stable one (following the ideas exposed in Section 4) should be used in each situation.

We are interested in the minimal requirements in a smooth shape for robust estimation of a projective transformation. It can be easily proved that a single concavity is sufficient. The idea is to set up a projective frame with one side on the bitangent, the other three sides tangent to the convex hull of the figure, and with both diagonals tangent to the concavity (Fig. 8).

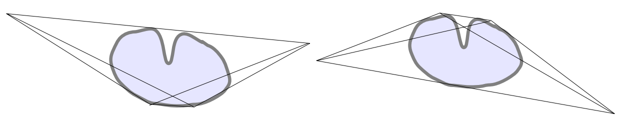


(a) (b) (c) (d)

Fig. 8. Projective frame from a single concavity. (a) and (b) are the contours of the two views of the lake in Fig. 1. (c) and (d) illustrate the dependence of the construction on the desired cross ratio of the intersections of the diagonal (c) = 0.36, (d)=0.01.

3.1 Existence and Uniqueness of the Construction

We outline an informal existence argument. Given the convex hull of the shape and the convex hull of the concavity we can set an 'initial', extremely distorted projective frame with diagonals 'including', but not touching, the concavity (Fig. 9.a), with two points extremely close and three sides nearly collinear. If the base extremes move closer to the shape, the diagonals will eventually touch the concavity, since we can always set up another extremely distorted frame intersecting the concavity (Fig. 9.b). Note that to achieve the desired double tangency the positions of the extremes are not independent from each other; there is a one-parameter family of solutions.



(a) (b)

Fig. 9. The diagonals of a projective frame can always be tangent to a concavity

The bitangent fixes one d.o.f. in the projective frame in addition to the previous four shown in Fig. 7 (left), and two more d.o.f.'s can be fixed by making the two diagonals of the projective frame tangent to the concavity. Uniqueness can be achieved by eliminating the remaining d.o.f. with a predetermined cross-ratio in the intersections of one diagonal and the convex hull (Fig. 8.c-d)

3.2 Algorithmic Complexity

In contrast with the two (or more) concavities case, where the projective reference frame can be directly constructed from the immediately available bitangents, working with a single concavity requires some search. We recommend the following algorithm. From an arbitrary starting point k in the bitangent (Fig. 10) we compute the tangent t to the concavity and the intersections a and b. The chosen cross-ratio fixes the opposite corner q in the frame[2]. From the tangents from q to the convex hull we obtain the intersections c and d.

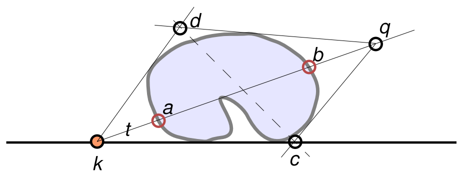
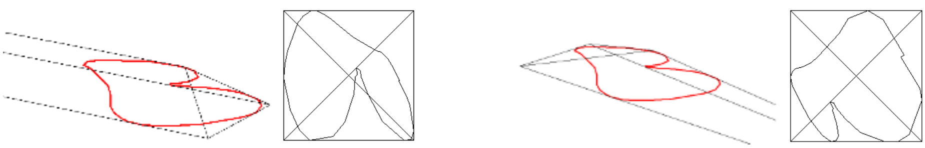


Fig. 10. Projective frame construction (see text)

Now we have only two possibilities: either the diagonal cd crosses the concavity, or not (the case shown in the figure). Given a k_{right} close enough to the left contact point of the bitangent (remember that the exact location of this point is not reliable) this diagonal will intersect the concavity (Fig. 11.a). Alternatively a k_{left} sufficiently far from the shape induces a diagonal that will not touch the concavity (Fig. 11.b). From the starting k_{left} and k_{right} positions we perform a binary search for the solution k^* with a cd diagonal tangent to the concavity (in practice, tangency can be acceptable if the diagonal intersects the convex-hull of the concavity in two points sufficiently close).



(a) k_{right} (b) k_{left}

Fig. 11. The two cases in the binary search of the concavity double tangency (see text)

In our experiments the projective frame is computed in a search process taking about 10 steps. Each tentative frame construction takes linear time with respect to polyline size and no polyline transformation, smoothing or preprocessing must be performed in the search, so the algorithm is extremely efficient. The overall computation time is negligible in relation to the image processing tasks required to extract the contours.

The construction becomes ill-conditioned when the contact points of the bitangent are very close (the concavity is nearly a hole) and when the concavity is too deep or too flat (three points in the reference become nearly collinear). In this paper we focus on constructions using a single concavity, even if the shapes have more than one, in order to evaluate the most adverse situation.

[2] We must check that q is in the correct side (the horizon is not 'crossed'), since some extreme k positions are incompatible with a frontal view.

3.3 Shape Similarity

A planar curve can be described by a continuous function $f : (0,1) \in \mathbb{R} \to \mathbb{C}$. A reasonable similarity measure for closed contours is the mean squared distance between 'homologous points': $d(f,g) = \int_0^1 (|f(t) - g(t)|^2 dt$. From the Parseval theorem this can be immediately computed in the frequency domain provided that the parameterization of both curves is consistent (normalization of the starting point involves a simple modification of the phase of the spectral coordinates). The desired Fourier coefficients of a closed polyline with arbitrarily spaced knots can be efficiently computed without need of resampling using the technique proposed in [20, Appendix]. The canonical version of a shape (projectively normalized by transforming the reference frame to the unit square) can be characterized by its low frequency coordinates. However, precise error alignment must be computed in the reference frame of the views.

4 Robustness Analysis

The proposed projective frame is built using only global properties of the shape. Local projective invariants, extremely sensitive to noise, are avoided. Therefore, it is expected that homography estimations based on it are robust against moderated amounts of noise. In this section we suggest a theoretical, rigorous approach to the study of the stability of the above construction and also describe a more practical stability assessment method used in our experiments.

For simplicity we quantify the level of noise in the imaging process (including acquisition, color thresholding or edge extraction and linking) by a single magnitude ϵ defined as the *maximum* distance from a true point in the 'ideal' contour and the corresponding 'corrupt' point (e.g., in certain cases ϵ could be related to pixel size). Therefore, the true shape lies inside a tolerance band around the observed contour. From this band we can compute the extreme constructions and report the *worst case* alignment situation for a given level of noise (Fig 12).

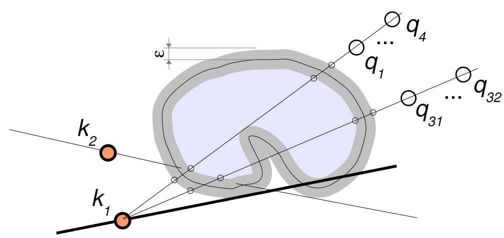


Fig. 12. Possible constructions induced by noise level ϵ (only a few points are shown)

An alternative, empirical approach is based on repeated computation of the projective frame from contour perturbations of at most size ϵ and report the distribution of alignment errors. Stability can be also assessed by alignment of the shape with a perturbed version of itself. Since we must go (loosely speaking) through the canonical frame and return, this kind of self alignment error is related to the quality of the shape for homography estimation.

A more practical stability measure can be directly derived from the own structure of the construction. The vertices of the projective frame are intersections of discrete tangent lines whose points of contact have error as large as ϵ (Fig. 13). Even if the rest of the construction is noiseless, a certain intersection x will have an uncertainty $\Delta x = S\epsilon$, where $S = \overline{px}/\overline{pc}$.

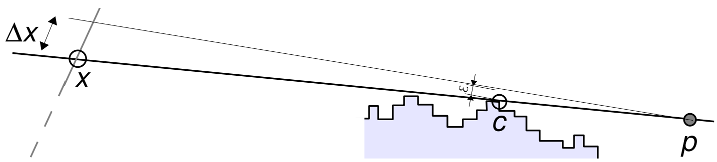


Fig. 13. Stability of a polygon tangent

The overall stability of the frame is in some sense dominated by the worst ingredient in the construction, so, for instance, an approximate unstability measure is the maximum 'error amplification' ratio S of all tangents.

5 Experiments

Fig. 14 shows the quality of the alignment of the lake contours, including the alignment error E (measured in normalized MSE distance $\times 1000$), and the simple unstability measure S ($\times 10$) of the constructions explained above. Observe that shape (a) is less stable ($S = 7.9$) than (b) ($S = 4.4$), as intuitively expected from the lengths and angles of the constructed frames. Even though the contours have been extracted with low precision and from completely unrelated sources, the proposed global procedure is still able to satisfactorily align both shapes directly from the raw available polylines. Note that alternative methods based on identification of homologous points or lines would require some kind of intelligent interpretation of the shape.

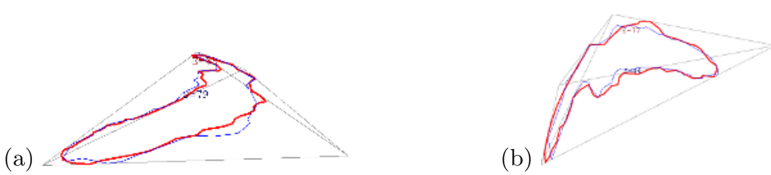


Fig. 14. Alignment of the lake shapes in Fig 1

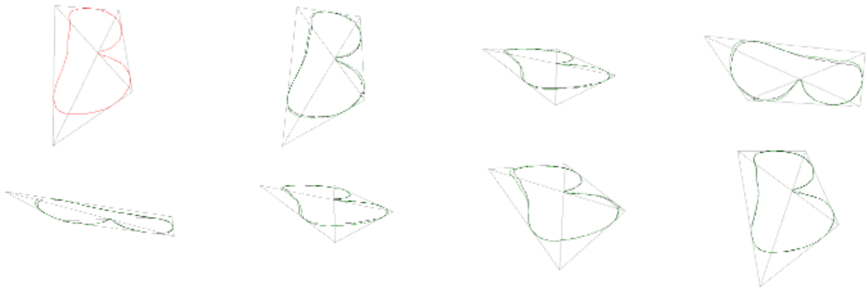


Fig. 15. Some frames in real-time alignment of a smooth shape (see text)

Fig. 15 shows real time alignment of a smooth, handwritten 'B' shape in a video sequence taken by a camera which moves freely in space. The first frame is the target and the rest are some illustrative views, most of them specifically selected with perturbations in the contour to demonstrate the robustness of the method. The full video sequence and additional demonstrations can be downloaded from the web page `http://ditec.um.es/contour`. Alignment is also acceptable on significantly reduced polylines (Fig. 16).

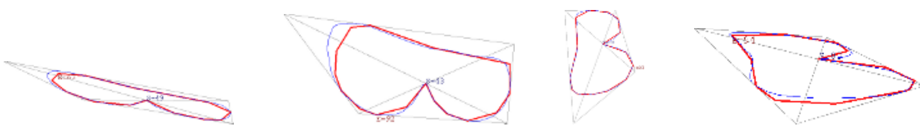


Fig. 16. Alignment on reduced polylines

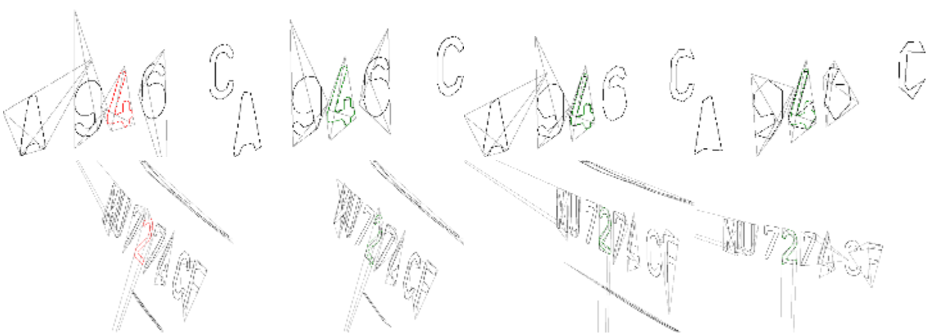


Fig. 17. Symbol recognition

Fig. 17 shows some examples of traffic plate symbol recognition for increasing noise levels, caused again by the tolerance in polyline reduction. (In this case an affine model is sufficient for shape recognition in views with small slant.)

Finally, Fig. 18 illustrates the estimation of camera pose [2] in a video sequence using the alignment homographies obtained from a smooth contour. We assume that the camera parameters are known.

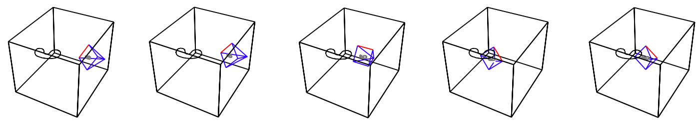


Fig. 18. Estimated 3D camera trajectory

6 Conclusions and Future Work

This paper proposes a novel projectively invariant representation of planar contours based on global convexity properties. We have shown that a canonical frame can be efficiently extracted from shapes with at least one appreciable concavity, using a remarkably simple geometric construction working from raw, irregularly sampled polylines. The stability of the reference frame has been formally studied and the maximum error amplification ratio has been proposed as a pragmatic measure of shape quality for projective alignment. Our experiments indicate that the homographies estimated by this method are surprisingly accurate even for considerable noise levels. In such extreme conditions alternative methods based on finding correspondences of local properties such as hot points, straight lines or conic approximations produce unacceptable results.

The method can be applied to image-model homography estimation, shape normalization and recognition, and even pose localization (given some knowledge of camera parameters). All these tasks can be performed in real time: the construction has linear algorithmic complexity with respect to the number of polyline knots, so the computational effort required by homography estimation is negligible in relation to the rest of low-level image processing stages.

This work can be extended in several directions. First, self-consistency tests must be implemented to avoid ill-conditioned configurations (for instance, contours with very small concavities). Alignments produced by extreme projective transformations should also be automatically detected and rejected. Second, a characterization of admissible occlusions (those which do not disturb the construction of the projective frame) would be very attractive for applications in cluttered environments. Finally, a theoretical model of alignment degradation should be rigorously developed in terms of noise level and some appropriate stability measure of the projective frame.

Acknowledgments

The authors would like to thank the anonymous reviewers for their useful suggestions. This work has been supported by the Spanish MCYT grants DPI2001-0469-C03-01 and TIC2003-08154-C06-03.

References

1. Pizlo, Z., Rosenfeld, A.: Recognition of planar shapes from perspective images using contour-based invariants. Computer Vision, Graphics, and Image Processing **56**(3) (1992) 330–350
2. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. 2 edn. Cambridge University Press (2004)
3. Faugeras, O., Luong, Q.T., Papadopoulo, T.: The Geometry of Multiple Images. MIT Press (2001)
4. Mundy, J., Zisserman, A.: Appendix – Projective geometry for machine vision. In: Geometric Invariances in Computer Vision, MIT Press (1992)
5. Rothwell, C.A., Zisserman, A., Forsyth, D.A., Mundy, J.L.: Canonical frames for planar object recognition. In: Proc. 2nd European Conference on Computer Vision, Santa Margherita Ligure, Italy, (1992) 757–772
6. Carlsson, S., Mohr, R., Moons, T., Morin, L., Rothwell, C., Diest, M.V., Gool, L.V., Veillon, F., Zissermann, A.: Semi-local projective invariants for the recognition of smooth plane curve. IJCV **19**(3) (1996) 211–236
7. Salden, A., Haar, B., Viergever, R.: Affine and projective differential geometric invariants of space curves. In: Baba Vemuri, ed., Geometric Methods in Computer Vision II, SPIE. (1993)
8. Zisserman, A., Blake, A., Rothwell, C., Van Gool, L., Van Diest, M.: Eliciting qualitative structure from image curve deformations. In: Proc. 4^{th} IEEE ICCV. (1993) 340–345
9. Weiss, I.: Noise resistant invariants of curves. IEEE PAMI **15**(9) (1993) 943–948
10. Lei, Z., Blane, M.M., Cooper, D.B.: 3L fitting of higher degree implicit polynomials. In: Proc. 3^{rd} IEEE WACV, Sarasota (USA) (1996)
11. Tarel, J., Civi, H., Cooper, D.B.: Pose estimation of free-form 3D objects without point matching using algebraic surface models. In: Proc. 1^{st} IEEE Workshop on Model-Based 3D Image Analysis, Mumbai (India) (1998) 13–21
12. Schmid, C., Zisserman, A.: The geometry and matching of curves in multiple views. In: Proc. 5^{th} ECCV, Freiburg (Germany) (1998) 394–409
13. Startchik, S., Milanese, R., Pun, T.: Projective and photometric invariant representation of planar disjoint shapes. Image and Vision Comp. **16**(9-10) (1998) 713–723
14. Chesi, G., Malis, E., Cipolla, R.: Collineation estimation from two unmatched views of an unknown planar contour for visual servoing. In: Proc. 10^{th} BMVC, Nottingham (UK) (1999)
15. Sato, J., Cipolla, R.: Extracting group transformations from image moments. Computer Vision and Image Understanding **73**(1) (1999) 29–42
16. Basri, R., Jacobs, D.: Projective alignment with regions. PAMI **23**(5) (2001) 519–527
17. Mendonca, P., Wong, K., Cipolla, R.: Camera pose estimation and reconstruction from image profiles under circular motion. In: Proc. 6^{th} ECCV, Dublin (Ireland) (2000) 864–877
18. Wong, K., , Cipolla, R.: Structure and motion from silhouettes. In: Proc. 8^{th} IEEE ICCV, Vancouver (Canada) (2001) 217–222
19. Cipolla, R., Giblin, P.: Visual Motion of Curves and Surfaces. Cambridge University Press (2000)
20. Arbter, K., Snyder, W., Burkhardt, H., Hirzinger, G.: Application of affine-invariant fourier descriptors to recognition of 3D objects. PAMI **12**(7) (1990) 640–647

21. Startchik, S., Milanese, R., Rauber, C., Pun, T.: Planar shape databases with affine invariant search. In: Proc. 1^{st} Int. Workshop on Image Databases and Multimedia Search, Amsterdam (Netherlands) (1996)
22. Vinther, S., Cipolla, R.: Object model acquisition and recognition using 3D affine invariants. In: Proc. 4^{th} BMVC, Guilford (UK) (1993) 369–378
23. Stolfi, J.: Oriented Projective Geometry. Academic Press, Boston (1991)
24. Melkman, A.: On-line construction of the convex hull of a simple polygon. Information Processing Letters **25** (1987) 11–12
25. López-de-Teruel, P.E., Ruiz, A., Fernández, L.: Geobot: A high level visual perception architecture for autonomous robots. In: Proc. 4^{th} IEEE ICVS, New York (USA) (2006)