

Recognizing Objects Using Scale Space Local Invariants

A. M. Bruckstein *

E. Rivlin †

I. Weiss ‡

Abstract

In this paper we discuss a new approach to invariant signatures for recognizing curves under viewing distortions and partial occlusion. The approach is intended to overcome the ill-posed problem of finding derivatives, on which local invariants usually depend. The basic idea is to use invariant finite differences, with a scale parameter that determines the size of the differencing interval. The scale parameter is allowed to vary so that we obtain a "scale space"-like invariant representation of the curve, with larger difference intervals corresponding to larger, coarser scales. In this new representation, each traditional local invariant is replaced by a scale-dependent range of invariants. Thus, instead of invariant signature curves we obtain invariant signature surfaces in a 3-D invariant "scale space".

1 Introduction

One of the major problems of object recognition is the fact that on one hand, an object can be seen from different points of view, producing different images. On the other hand, we would like to store only one image in a data base and match any other image of the object to it, regardless of the point of view. A good way to overcome this problem is to use viewpoint invariants, namely descriptors of the the shape that are independent of the point of view, and use them for matching.

The subject of viewpoint invariants in vision has developed rapidly in recent years. A simple projective, or viewpoint, invariant, namely the cross ratio of four points on a line, was introduced in vision by Duda and Hart [1]. However, its domain of applicability was very limited. More general invariants were studied in the nineteenth century, and were introduced in the field of computer vision by Weiss

[2]. They are of two main types: 1) Algebraic invariants. These are based on a global description of the shapes by algebraic entities such as lines, conics and polynomials. 2) Differential invariants. These are based on describing the shape by arbitrary differentiable functions.

These methods have been applied to various vision problems. The algebraic approach was used by Forsyth et al. [3] and Taubin Taubin, while differential invariants were used by Weiss [2] and Bruckstein and Netravali [4]. Both methods proved to have advantages and disadvantages. The algebraic method, while simple and easy to implement, is quite limited in the kinds of shapes that it can handle because most shapes are not representable by simple low order polynomials. The differential method is more general because it can handle arbitrary curves, but it relies on the use of local information such as derivatives (of quite high orders).

This situation has led to the introduction of various kinds of intermediate, or hybrid methods, that try to combine the advantages of the algebraic and differential methods, and hopefully not their disadvantages. Van Gool et al. [5], Brill et al. [6], and others introduced invariants that contain both derivatives and reference points. Each reference point reduces the number of derivatives that one needs in order to obtain invariants. In [10] a "canonical" coordinate system was used without curve parameterization to obtain the same goal. This resulted in fewer derivatives and in the capability of using feature lines in addition to points. However, in all these methods, the correspondence must be established between the reference points of the two images that are being matched. Finding the correspondence is a very difficult problem that requires searches in high dimensional spaces, and we need a method that avoids this.

In this paper we reduce the number of derivatives by using a scale space approach. It is well known (Koenderink [7]; Witkin [8]) that such an approach can turn the ill-posed problem of finding derivatives into a well-posed one. The scale space has to be invariant so we cannot use simple Gaussian-like smoothing. Instead, we rely on some reference points as a function of the given curve and a variable scale parameter. These reference points are not assumed to be readily available in the image, as in previous methods, but are determined from the curve in an invariant way. Thus, no correspondence is needed. Using low-order derivatives

*Department of Computer Science, Israel Institute of Technology, Technion, Haifa, Israel. Fax: 972-4-294353. Email: freddy@cs.technion.ac.il

†Department of Computer Science, Israel Institute of Technology, Technion, Haifa, Israel. Fax: 972-4-294353. Email: ehudr@cs.technion.ac.il

‡Computer Vision Laboratory, Center for Automation Research University of Maryland, College Park, MD 20742-3275, Fax: 301-314-9115 Email: weiss@cfar.umd.edu

and our variable reference points, we build invariant scale space representations of the given curves.

There are various ways to derive invariants in accordance with the above scheme. Here we extend a method originally introduced by Bruckstein et al. [4]. It consists of defining an invariant arclength (using the lowest possible order of derivatives in given schemes), and then defining invariant finite differences using this arclength. These differences replace the higher order derivative in the traditional invariants. The differences are not necessarily small and do not tend to zero. Rather, their variable size creates the “scale space”.

We briefly describe here an illustrative example of the method. Given a curve, we want to find invariants at each point of the curve so that we can obtain a local invariant signature. With Euclidean invariance in mind, we can plot the curvature vs. the arclength τ to obtain a Euclidean invariant signature. Invariant signature plots of two curves are then compared to detect matches rather than the curves themselves. This is an example of a local method, in which no correspondence between points is needed. However, curvature involves a second derivative which we wish to avoid.

In our new method, the second derivative is replaced by a finite difference. We start from a point $P(\tau)$ on the curve, and we want to find invariants there. We choose an interval size $\Delta\tau$ and find two points on the curve, $P(\tau + \Delta\tau)$, $P(\tau - \Delta\tau)$, located at distances $+\Delta\tau$ and $-\Delta\tau$ (measured on the curve) from the point $P(\tau)$ at which we want to calculate the invariants. Given these three points, we can calculate any Euclidean invariant involving them, such as the area $A(\tau)$ of the triangle formed by them. $A(\tau)$ is then a new type of invariant signature. This is much more robust than a derivative, if $\Delta\tau$ is not too small. In this way we reduce the number of derivatives needed without needing any fixed reference points or their correspondence. The scale parameter $\Delta\tau$ can now be varied to obtain a whole range of scale dependent invariants.

In summary, the semi-local, or finite difference method elaborated upon in Bruckstein, Holt et al. [9] is extended here as follows. First we consider more general transformations such as similarity, affine, or even projective viewing distortion. We use similarity, affine or projective invariant arclength to reparametrize the curve, exploiting all the information available. We then let the differencing interval size or sizes be free parameters rather than setting them in advance. In this way we obtain whole ranges of invariants at each point rather than single values. The signature functions for the curves then become signature vectors or even continua of values, i.e., surfaces or hypersurfaces. Matching them will be slightly more complicated but will certainly be robust because it will be less sensitive to peculiarities that may exist at some fixed pre-set value of the locality (scale) parameters.

2 Theory of Scale Dependent Local Invariants

Here we describe in detail the basic ideas of the semi-local method. Its main advantage over the global method is its ability to deal with partially occluded shape. We deal here with planar curves, such as boundaries of planar objects.

To obtain an invariant representation of a curve, we associate with each point of the curve a set of invariants. The collection of independent invariants from all points is the invariant “signature”. This approach maps the problem into a problem of detecting partial matches between the signatures of the “library” of possible objects and the signature functions extracted from the (composite) objects appearing in the scene to be analyzed.

We treat here two variants of such signatures. In the first a signature with an arclength as an independent variable. We first derive an invariant arclength τ and use it to reparametrize the curve. After that another invariant I is determined at each point (for example, curvature in the Euclidean case), and we represent the signature as a function $I(\tau)$. In the second a signature with two independent absolute invariants as coordinates. Here we find two local invariants I_1, I_2 at each curve point, and then plot I_2 against I_1 . The functions I_1, I_2 may or may not be represented as functions of an invariant arclength τ , but they are local quantities in any case. In both cases, recognition is based on detecting portions of invariant traces in the “transform plane”.

The simple Euclidean example will again be used to clarify the above discussion. Suppose we wish to detect the presence of partially occluded planar objects whose instances may undergo planar rotations and translations (i.e. Euclidean transformations). Here the well-known invariant signature approach describes object boundaries via curvature versus arclength functions, invariant under Euclidean transformation, and recognition is possible by partial matching. This method of finding a signature is based on using an invariant metric on the curve (the arclength) and on finding a differential invariant at each point on the curve (the curvature).

The second approach above was used in [10], without a scale space. No curve parametrization was used there. The method is based on the fact that a general curve has two independent invariants at each point, and these can determine the curve uniquely up to the relevant transformation. Here we describe various ways of using a the method in scale space. We deal with the case of Euclidean invariance. One way to proceed is to use a Euclidean arclength t . Associate with points on the arbitrarily parametrized curve $P(t) = [x(t), y(t)]$ two numbers $I_1(t)$ and $I_2(t)$ invariant under Euclidean transformation. Here, $I_1(t)$ could be the curvature at $P(t)$ and $I_2(t)$ could be the area of the triangle

formed by the points $\{P(t), P(t - t_b), P(t + t_f)\}$ where $P(t - t_b)$ is located at an *arclength* distance of t_b (chosen a priori “before” $P(t)$ and $P(t + t_f)$ at a distance of t_f “following” $P(t)$ in the traversal of the curve (see Figure 1). A more appealing way to use the second method is to avoid the curve parameter altogether. We can avoid it by a variety of methods. For example, we can define the first invariant at $P(t)$ to be the area between the curve and a parallel to the tangent at $P(t)$ at a distance of D (set beforehand) toward the center of the osculating circle.

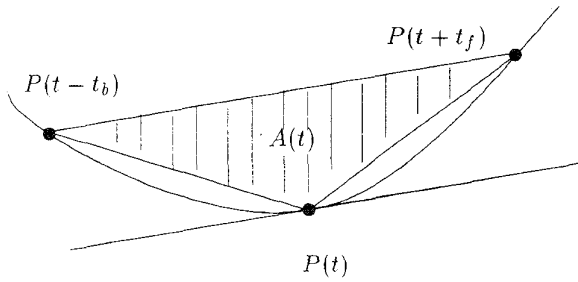


Figure 1. $P(t) \rightarrow [I_1(t) = \frac{1}{R(t)}, I_2(t) = A(t)]$

To summarize our message: The first, metric-based approach to finding a signature, calls for the arclength reparametrization of the curve, i.e., $P(t) \rightarrow P(\tau)$, and the association of one invariant quantity with each point of the reparametrized curve (in the above example—the curvature). The second, invariant coordinate approach, associates different invariant quantities with each point of the curve without necessarily referring to a curve parametrization.

Note that both approaches are based on our ability to analyze the neighborhood of a point on a planar curve and calculate some quantity that remains the same when we consider the *image of the point* and the *image of its neighborhood* under the viewing transformation.

In this paper we concentrate on the first method discussed above, i.e. we assume that we can always determine an invariant metric on the curve. With this metric, moving to the left and right along the curve from a point $P(\tau)$ to points at “distances” $\pm\Delta\tau_1, \pm\Delta\tau_2, \pm\Delta\tau_3, \dots$, etc., is a well-defined process. This process can be used to generate point sets anchored at $P(\tau)$ that are invariant under the distorting viewing transformation. Based on these point sets, we are able to use the global invariants of the viewing transformation to calculate a wide variety of invariants.

More importantly, notice that the point sets are parametrized by the sequence of positive numbers $\Delta\tau_1 < \Delta\tau_2 < \Delta\tau_3 < \dots$ and hence the invariant quantities that we generate are likewise parametrized. We can use this freedom to associate with each point $P(\tau)$ on the curve a

whole range of invariants rather than a single one. Hence we can define *multi-valued* or *parametrized* signatures (or coordinates in the second approach discussed above). These have the potential of enabling more robust matching in the presence of noise and other disturbances.

To illustrate this let us again consider the Euclidean case. Once the curve $P(t)$ is reparametrized to $P(\tau)$ where τ is Euclidean arclength we can proceed as follows: At each point $P(\tau)$, consider the point set $\{P(\tau - \Delta\tau), P(\tau), P(\tau + \Delta\tau)\}$ and compute the radius of the circle passing through these points, denoting it by $R(\tau, \Delta\tau)$. Clearly as $\Delta\tau \rightarrow 0$, $R(\tau, \Delta\tau) \rightarrow \frac{1}{k(\tau)}$ where $k(\tau)$ is the curvature. However, we can use the whole range of values of $\Delta\tau$ from $\Delta\tau = 0$ to some $(\Delta\tau)_{\max}$ to associate with $P(\tau)$ a multi-valued curvature function of the form $k(\tau, \Delta\tau) = \frac{1}{R(\tau, \Delta\tau)}$. At $\Delta\tau = 0$ we obtain $k(\tau)$, but $k(\tau, \Delta\tau)$ clearly carries more information on the local behavior of the curves around $P(\tau)$ than $k(\tau)$, in the neighborhood of any value of τ .

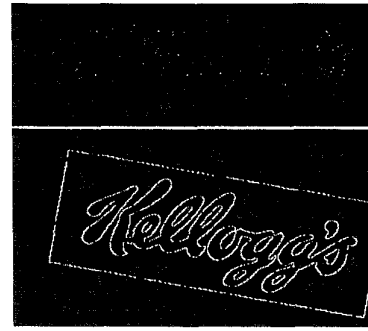


Figure 2. A logo before and after transformation (scaling and rotation). The logo was processed as five different curves.

Furthermore, we can use other Euclidean invariant quantities, like the areas of the triangles $\{P(\tau), P(\tau - \Delta\tau), P(\tau + \Delta\tau)\}$, i.e., $A(\tau, \Delta\tau) = \text{Area}\{P(\tau), P(\tau - \Delta\tau), P(\tau + \Delta\tau)\}$, the angles $\varphi(\tau, \Delta\tau) = \angle P(\tau - \Delta\tau)P(\tau)P(\tau + \Delta\tau)$, etc. All these are valid “generalized parametrized signature” functions that can be associated with a planar curve. (There are clearly relationships between the various quantities, but this will not concern us here). In case we need to recognize occluded planar shapes under a Euclidean viewing transformation these generalized signatures will enable us to perform more robust partial matching for detection.

Note that the same approach can also be used in conjunction with the invariant coordinate method. One of the invariant quantities can be chosen as the independent variable $\tau = I_1(t)$, and the other can be a parametrized continuum of values $I_2(\tau, \Delta\tau)$.

Note also the important point that we do *not* necessarily

advocate the computation of the limit values for $\Delta\tau \rightarrow 0$. If $\Delta\tau$ takes only a finite set of positive values we base our invariants on a form of *finite differences* in the invariant metric, rather than on the differential behavior of the curve about $P(\tau)$.

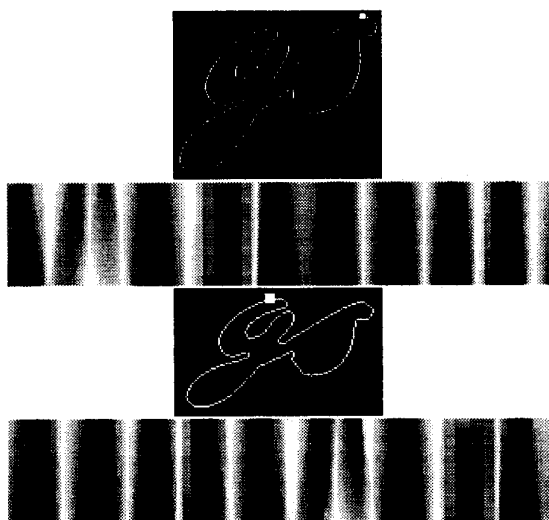


Figure 3. The last two letters of the logo (*gs*) constitute a curve to be processed. The multi-valued signatures are presented below. Shifting the top signature to the right by approximately half the strip's length will achieve a match.

3 Experiments: Invariant Scale Space Signatures

We present a series of experiments to illustrate the above-outlined theory. The similarity invariant arclength parameter is given in this case by

$$d\tau = \left| \frac{x'(t)y''(t) - x''(t)y'(t)}{(x'(t))^2 + (y'(t))^2} \right| dt$$

After the curve is reparametrized by the invariant arclength we can call upon several types of scale-dependent similarity invariants. In this example (and in the ones that follow), we plot the angle $[P(\tau - \Delta\tau) P(\tau) P(\tau + \Delta\tau)] = \varphi(\tau, \Delta\tau)$ as a function of τ . However, a wealth of other possibilities are available. We could also compute various length or area ratios that are also known to be similarity invariants.

In our experiments each image contains 20 different signatures for 20 different parameter values. For each signature different $\Delta\tau$ s were used. For a constant Y value one gets

single-valued signatures for the curve. The grey level encodes the similarity invariant for a particular arclength and parameter value. The full display represents an “invariant signature surface”. For each curve the starting position is marked by a white square. Due to the different starting position one multi-valued signature is shifted relative to the other. To check for a match between two signatures one should match one multi-valued signature to the other while shifting it in a cyclic manner. This is done automatically, scoring each match, trying to solve for the maximum. In this example one can see that a match is achieved when one of the signatures is shifted. The symmetric nature of the curve is evident in the signature structure.

For the following experiments we used images of different logos. Each logo is processed to an edge image. Curves are processed by length, and a B-spline is matched to each of the curves. Using the B-spline derivatives are computed, and invariant arclength is obtained. The figures show the results of the processing on two images of logos going under similarity transformation. In Figure 2, a familiar logo was processed and mapped into five different curves. The scale-space signatures for two curves out of the five are shown in Figure 3. In both cases a good match is achieved modulo a (circular) shift in the invariant arclength.

In the following we show results of experiment that handle occlusion. The processing stage is similar. When we produce signatures for open curves using different parameters we have different domain for each parameter. Hence, extracting multi-valued signatures from occluded curve forces us to further reduce the common domain. As a result we are restricted with the amount of occlusion we can handle without difficulty. Still occlusion of 30% of the shape can be handled without any problem. The automatic matching is done in the same manner as before, comparing the multi-valued signature in each step in a simple cyclic move. The best score achieved is the result of the comparison. In this experiment each multi-valued signature contains only five different signatures for five different parameter values. Figure 4 show a logo under occlusion. The multi-valued signatures of the occluded curves are presented below the image of the logo. One can see that a good match is obtained.

For the affine case, as noted before, the invariant arclength parameter can be obtained from any arbitrary parameter s by

$$d\tau = |(x_t y_{tt} - x_{tt} y_t)^{1/3}| dt$$

with the subscripts denoting derivatives with respect to t . This expression is invariant except for a factor equal to the determinant of the transformation. To make the arclength invariant to this factor too, we normalize the expression above by the total arclength $\int d\tau$ of the curve segment we deal with. After the curve is reparametrized by the invariant arclength we can call upon several types of affine invariants. In the example we plot areas ratio against the invariant ar-

length parameter. In this experiment too, each multi-valued signature contains five different signatures for five different parameter values.

4 Discussion

We have developed a way of improving the reliability of object recognition by the method of local invariants. The advantage of local invariants relative to global ones is their ability to handle occlusion. The difficulty in using them lies in the need to use derivatives. Derivatives are not very robust to noise, and even in the noiseless case they can depend on the scale at which we look at the image, namely the degree of smoothing. We have proposed solving this problem by looking at the shape at many scales rather than trying to choose one particular scale factor, which is not invariant.

Instead of derivatives, we use a finite difference method in an invariant form. The differences depend on a scale factor which we allow to vary continuously, thereby obtaining a description of the shape in an invariant scale space. Scale space methods have been extensively used, but mostly not in an invariant way. The treatment here is quite general; several forms of difference-based invariants have been treated here for projective, affine and similarity transformations. We have shown experimentally that the method can easily recognize various complicated shapes.

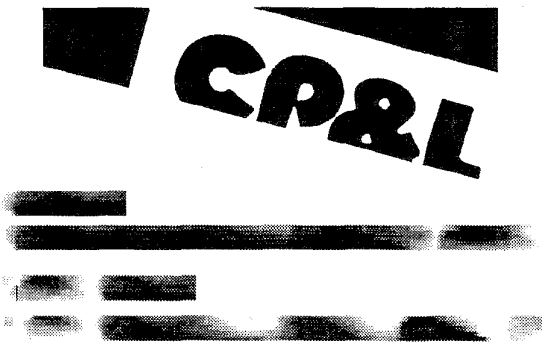


Figure 4. The logo was processed as four different curves. The multi-valued signature for the letter P under occlusion, and below it the signature for the complete curve. The signatures below belong to the sign &. The presented relative position of the two signatures is the one which gives the best match.

References

[1] Duda, R.O. and Hart, P.E., *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.

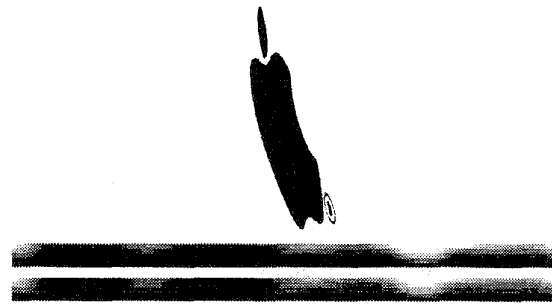


Figure 5. An input logo under affine transformation and the multi-valued signature for the Apple logo. On top the signature for the database logo, below the signature for the logo after the transformation. The presented relative position of the two signatures is the one which gives the best match.

[2] Weiss, I., Projective Invariants of Shapes, *Proc. DARPA Image Understanding Workshop*, Cambridge, MA, 1125–1134, 1988.

[3] Forsyth, D., Mundy, J.L., Zisserman, A., and Brown, C.M., Projectively Invariant Representations using Implicit Algebraic Curves, *Image and Vision Computing* 8, 130–136, 1990.

[4] Bruckstein, A.M. and Netravali, A.N., On Differential Invariants of Planar Curves and the Recognition of Partially Occluded Planar Shapes, AT&T Technical Memo, July 1990.

[5] Van Gool, L., Kempenaers, P., and Oosterlinck, A., Recognition and Semi-Differential Invariants, *Proc. CVPR*, 454–460, 1991.

[6] Barrett, E., Payton, P., Haag, N., and Brill, M., General Methods for Determining Projective Invariants in Imagery, *CVGIP:IU* 53, 45–65, 1991.

[7] Koenderink, J.J. and Van Doorn, A.J., Affine Invariants from Motion, *J. Opt. Soc. Am. A*, 377–385, 1991.

[8] Witkin, A.P., Scale-Space Filtering, *Proc. IJCAI*, 1019–1022, 1983.

[9] Bruckstein, A., Holt, J., Netravali, A.N., and Richardson, T.J., Invariant Signatures for Planar Shape Recognition under Partial Occlusion, *CVGIP:IU* 58, 49–65, 1993.

[10] Rivlin, E. and Weiss, I., Local Invariants for Recognition, *IEEE-PAMI* 17, 226–238, 1995.