

Highly Discriminative Invariant Features for Image Matching *

Ronald Alferez and Yuan-Fang Wang

Department of Computer Science

University of California

Santa Barbara, CA 93106

E-mail: {ronald,yfwang}@cs.ucsb.edu

Abstract. In this paper, we present novel image-derived, invariant features that accurately capture both the geometric and color properties of an imaged object. These features can distinguish between objects that have the same general appearance (e.g., different kinds of fish), in addition to the typical task of distinguishing objects from different classes (e.g. fish vs. airplanes). Furthermore, these image features are insensitive to changes in an object's appearance due to rigid-body motion, affine shape deformation, changes of parameterization, perspective distortion, view point change and changes in scene illumination. The new features are readily applicable to searching large image databases for specific images. We present experimental results to demonstrate the validity of the approach, which is robust and tolerant to noise.

1 Introduction

The advent of high-speed networks and inexpensive storage devices makes the construction of large image databases feasible. More and more images are now stored in electronic archives. In line with this, however, is the need for tools to help the user browse and retrieve database images efficiently and effectively.

Most existing image indexing and retrieval systems, such as Virage [4], QBIC [5], and Photobook [6], are able to do *between-classes* retrieval. That is, they can distinguish between images of *different* classes. For example, an image of a fish as a query retrieves a list of images in the database containing an image similar to a fish (the query and the generated results are classified as belonging to the same class of objects). Images that belong to other classes, such as airplanes, are appropriately excluded from the list. However, these systems do not allow the user to retrieve images that are more specific. In other words, they are unable to perform *within-a-class* retrieval. For example, the user may want to retrieve all images of rainbow trouts (characterized by the number and location of fins, and by the color of their body). Current systems will likely fail with this query, generating lists of images containing various species of fish. The problem is that a rainbow trout appears very similar to other species of fish, and the features adopted by current systems are not descriptive enough to handle this type of scenario. Hence, there is a need for a system that enables *within-a-class* retrieval, which discriminates between images within the same class of objects. In addition, environmental changes such as an object's pose and lighting should be not be a factor in measuring similarity.

To perform *within-a-class* retrieval in image databases, the system should be able to discriminate between imaged objects that have very similar appearance. The key to building such a system is in designing powerful, highly discriminative

* Supported in part by a grant from the National Science Foundation, IRI 94-11330

image features that can discriminate small variations among objects. These variations, however, should not include changes that are not intrinsic to an object, so that an object that is stretched, for example, should not be distinguished from its original form. Many digital library applications will find *within-a-class* retrieval particularly useful. Potential scenarios include searching for fish in an aquarium database, leaves and flowers in a botanical image database, and logos in a catalog. Despite the similar appearance of objects within each of these databases, and despite possible changes in pose and scene illumination, our new image features should be able to discriminate between different imaged objects within a database, while correctly matching the same ones.

Our contribution is in developing novel image-derived features that enable *both between-classes and within-a-class retrievals*. Not only do the new features discriminate between imaged objects that look very different, they can also distinguish between imaged objects with very similar appearance. Furthermore, these image features are insensitive to environmental changes such as rigid-body motion, affine shape deformation, changes of parameterization, perspective distortion, view point change and changes in scene illumination. These image features can be applied to image indexing, search and retrieval for large image databases, where high accuracy and environmental insensitivity is an issue.

Although segmentation (contour extraction) is not addressed, our strategy still has many practical applications, particularly when there is absolute control of the image database (*e.g.*, when the database is a collection of imaged objects photographed with an uncluttered background, such as catalogs), and the object of interest in the query image is pinpointed (or drawn) by a human.

We propose invariant features that capture only the essential traits of an image, forming a compact and intrinsic description of an imaged object. Environmental factors such as pose and illumination are ignored. Hence, it is more efficient than, say, aspect-based approaches where multiple aspects of the same model have to be remembered. The new invariant features analyze the shape of the object's contour as well as the color characteristics of the enclosed area. The analysis involves projecting the shape or color information onto one of many basis functions of finite, local support (*e.g.*, wavelets, short-time Fourier analysis, and splines). Invariance of the descriptors is achieved by incorporating the projection coefficients into formulations that cancel out many environmental factors. The invariant features produced by the new framework are insensitive to rigid motion, affine shape deformation, changes of parameterization and scene illumination, and/or perspective distortion. Furthermore, they enable a quasi-localized, hierarchical shape and color analysis, which allows for the examination of information at multiple resolution scales. The result is an invariant framework which is more flexible and tolerant to a relatively large degree of noise.

Excellent reviews on invariants are presented in [7, 8].

2 Technical Rationale

We will illustrate the design of invariant image features using a specific scenario where invariants for curves are sought. For shape invariants, these directly apply to the silhouette (contour) of imaged objects in a database. For illumination invariants, the same technique applies by linearizing internal regions by a characteristic sampling curve and computing invariant color signatures along the characteristic curve. In both cases, the invariant signatures produced can be

examined at different resolution scales, making the invariant features both flexible and noise tolerant. The particular basis functions we use in the illustration are the wavelet bases and spline functions. However, the same framework can be easily extended to other bases and to 3D surfaces.

Affine Invariant Parameterization We first look at the problem of point correspondence when attempting to match two curves (or contours) under an affine transformation. For each point selected from one curve, the corresponding point on the other curve has to be properly identified. In defining parameterized curves $\mathbf{c}(t) = [x(t), y(t)]^T$, the traditional arc length parameter, t , is not suitable because it does not transform linearly (or it is not invariant) under an affine transformation. Two parameterizations which do, are described in [2]: (1) The *affine arc length*, is defined as: $\tau = \int_a^b \sqrt[3]{\dot{x}\ddot{y} - \ddot{x}y} dt$ where \dot{x}, \dot{y} are the first and \ddot{x}, \ddot{y} are the second derivatives with respect to any parameter t (possibly the intrinsic arc length); and (2) the *enclosed area parameter*, $\sigma = \frac{1}{2} \int_a^b |xy - yx| dt$, which is the area of the triangular region enclosed by the two line segments from the centroid to two contour points a and b .

Seemingly, a common origin and traversal direction on the contour must also be established. However, it can be easily shown that a difference of starting points is just a phase-shift between the invariant signatures of two contours. Similarly, two contours parameterized in opposing directions are just flipped, mirror images of each other. Hence, a match can be chosen that maximizes the cross-correlation between the two signatures. This, together with the use of an affine invariant parameterization, implies that *no point correspondence is required* when computing the affine invariants of an object's contour.

Rigid Motion and Affine Transform Consider a 2D curve, $\mathbf{c}(t) = [x(t), y(t)]^T$ where t denotes a parameterization which is invariant under affine transform, and its expansion onto the wavelet basis $\psi_{a,b} = \frac{1}{\sqrt{a}}g(\frac{t-b}{a})$ [3] as $\mathbf{u}_{a,b} = \int \mathbf{c}\psi_{a,b}dt$. If the curve is allowed a general affine transform, we have: $\mathbf{c}'(t) = \mathbf{m}\mathbf{c}(\pm t + t_0) + \mathbf{t}$ where \mathbf{m} is any nonsingular 2×2 matrix, \mathbf{t} is the translational motion, t_0 represents a change of the origin in traversal, and \pm represents the possibility of traversing the curve either counterclockwise or clockwise². It follows that:

$$\begin{aligned} \mathbf{u}'_{a,b} &= \int \mathbf{c}'\psi_{a,b}dt &&= \int (\mathbf{m}\mathbf{c}(\pm t + t_0) + \mathbf{t})\psi_{a,b}dt \\ &= \mathbf{m} \int \mathbf{c}(t') \frac{1}{\sqrt{a}}g(\frac{\mp(t'-t_0)-b}{a})dt' + \int \mathbf{t}\psi_{a,b}dt &&= \mathbf{m} \int \mathbf{c}(t') \frac{1}{\sqrt{a}}g(\frac{t'-(\pm b+t_0)}{a})dt' \\ &= \mathbf{m} \int \mathbf{c}(t')\psi(t')_{a,\pm b+t_0} dt' &&= \mathbf{m}\mathbf{u}_{a,\pm b+t_0} . \end{aligned} \quad (1)$$

Note that we use the wavelet property $\int \psi_{a,b}dt = 0$ to simplify the second term in Eq. 1. If \mathbf{m} represents a rotation (or the affine transform is a rigid motion of a translation plus a rotation), it is easily seen that an **invariant expression** (this is just one of many possibilities) can be derived using the ratio expression

$$\frac{|\mathbf{u}'_{a,b}|}{|\mathbf{u}'_{c,d}|} = \frac{|\mathbf{m}\mathbf{u}_{a,\pm b+t_0}|}{|\mathbf{m}\mathbf{u}_{c,\pm d+t_0}|} = \frac{|\mathbf{u}_{a,\pm b+t_0}|}{|\mathbf{u}_{c,\pm d+t_0}|} . \quad (2)$$

The wavelet coefficients $\mathbf{u}'_{a,b}$ and $\mathbf{u}_{a,\pm b+t_0}$ are functions of the scale a and the displacements b and $\pm b + t_0$. If we fix the scale a , by taking the same number of sample points in each curve, we can construct expressions based on correlation

² In the implementation, the parameter is computed modularly over a closed contour.

coefficients to cancel out the effect of a different traversal starting point (t_0) and direction ($\pm t$). Let us define the **invariant signature** of an object, $f_a(x)$, as:

$$f_a(x) = \frac{|\mathbf{u}_{a,x}|}{|\mathbf{u}_{a,x+x_0}|} \quad \text{and} \quad f'_a(x) = \frac{|\mathbf{u}'_{a,x}|}{|\mathbf{u}'_{a,x+x_0}|} = \frac{|\mathbf{u}_{a,\pm x+t_0}|}{|\mathbf{u}_{a,\pm(x+x_0)+t_0}|}, \quad (3)$$

where x_0 represents a constant value separating the two indices. Then one can easily verify that when the direction of traversal is the same for both contours, $f'_a(x) = \frac{|\mathbf{u}_{a,x+t_0}|}{|\mathbf{u}_{a,x+x_0+t_0}|} = f_a(x+t_0)$. If the directions are opposite, then

$$f'_a(x) = \frac{|\mathbf{u}_{a,-x+t_0}|}{|\mathbf{u}_{a,-x-x_0+t_0}|} = \frac{1}{f_a(-x-x_0+t_0)}. \quad \text{As the correlation coefficient of two signals is defined as}$$

$$R_{f(x)g(x)}(\tau) = \frac{\int f(x)g(x+\tau)dx}{\|f\| \cdot \|g\|}.$$

We define the **invariant (similarity) measure** $I_a(f, f')$ between two objects as

$$I_a(f, f') = \max_{\tau, \tau'} \{R_{f_a(x)f'_a(x)}(\tau), R_{f_a(x)\frac{1}{f'_a(-x)}}(\tau')\}. \quad (4)$$

It can be shown [1] that the invariant measure in Eq. 4 attains the maximum of 1 if two objects are identical, but differ in position, orientation, and/or scale.

Other invariant features may still be derived where the same technique can be employed to measure similarity, making it independent of the parameterization used. For simplicity, we only show the invariant expressions from this point on.

It is known that the area of the triangle formed by any three $\mathbf{u}_{a,b}$ changes linearly in an affine transform [7]. Hence, we have the following invariants³:

$$\frac{\begin{vmatrix} \mathbf{u}'_{a,b} & \mathbf{u}'_{c,d} & \mathbf{u}'_{e,f} \\ 1 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} \mathbf{u}'_{g,h} & \mathbf{u}'_{i,j} & \mathbf{u}'_{k,l} \\ 1 & 1 & 1 \end{vmatrix}} = \frac{\begin{vmatrix} \mathbf{u}_{a,\pm b+t_0} & \mathbf{u}_{c,\pm d+t_0} & \mathbf{u}_{e,\pm f+t_0} \\ 1 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} \mathbf{u}_{g,\pm h+t_0} & \mathbf{u}_{i,\pm j+t_0} & \mathbf{u}_{k,\pm l+t_0} \\ 1 & 1 & 1 \end{vmatrix}}. \quad (5)$$

Perspective Transform Allowing an arbitrary view point and large perspective distortion makes the problem much harder as the projection is a non-linear process, involving a division in computing 2D coordinates. Extending the curve to 3D makes it even more difficult. A simplified model is possible, using a parallel or quasi-perspective (affine) model, but this holds only to a certain degree under a small perspective distortion. We provide a more rigorous treatment of perspective invariants. The projection process can be linearized using a tool which is well-known in computer graphics, the *rational* form of a basis function.

We will use NURBS (Non-Uniform Rational B-Spline) for illustration. The rational form of a b-spline function in 2D (3D) is the projection of a non-rational b-spline function in 3D (4D). Specifically, let $\mathbf{C}(t) = [X(t), Y(t), Z(t)]^T = \sum_i \mathbf{P}_i N_{i,k}(t)$ be a non-rational curve in 3D where \mathbf{P}_i 's are its control vertices and $N_{i,k}(t)$ are the non-rational spline basis. Its projection in 2D will be:

$$\mathbf{c}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} \frac{X(t)}{Z(t)} \\ \frac{Y(t)}{Z(t)} \end{bmatrix} = \sum_i \mathbf{p}_i R_{i,k}(t), \quad \text{where } R_{i,k}(t) = \frac{Z_i N_{i,k}(t)}{\sum_j Z_j N_{j,k}(t)}, \quad (6)$$

and \mathbf{p}_i 's are the projected control vertices in 2D, and $R_{i,k}$ are the rational bases.

We can now formulate the problem of finding perspective invariants as a curve fitting problem. Intuitively, if a 2D curve results from the projection of a 3D curve, then it should be possible to interpolate the observed 2D curve

³ Some may require a smaller number of coefficients. For example, for wavelet bases where $\int \psi_{a,b} dt = 0$, Eq. 5 can be simplified where only four coefficients are used.

using the projected control vertices and the *rational* spline bases and obtain a good fit. If that is not the case, then the curve probably does not come from the projection of the particular 3D curve. Hence, the error in curve fitting is a measure of invariance. (Ideally, the error should be zero.) Perspective projection produces:

$$\mathbf{p}'_i = \begin{bmatrix} \frac{X'_i}{Z'_i} \\ \frac{Y'_i}{Z'_i} \end{bmatrix} = \begin{bmatrix} \frac{r_{11}X_i + r_{12}Y_i + r_{13}Z_i + T_x}{r_{31}X_i + r_{32}Y_i + r_{33}Z_i + T_z} \\ \frac{r_{21}X_i + r_{22}Y_i + r_{23}Z_i + T_y}{r_{31}X_i + r_{32}Y_i + r_{33}Z_i + T_z} \end{bmatrix} \quad (7)$$

$$R'_{i,k} = \frac{(r_{31}X_i + r_{32}Y_i + r_{33}Z_i + T_z)N_{i,k}(t)}{\sum_j (r_{31}X_j + r_{32}Y_j + r_{33}Z_j + T_z)N_{j,k}(t)}. \quad (8)$$

where r_{ij} 's and T_i 's are the rotation and translation parameters, respectively. Image invariant defined by the goodness of fitting is $I = \int_t (\mathbf{d}(t) - \sum_i \mathbf{p}'_i R'_{i,k}(t))^2$, where $\mathbf{d}(t)$ denotes the distorted image curve. Note that in Eq. 6, the shape of a 2D curve is determined by the *projected control vertices* and the *rational* spline bases, *both* of which are unknown. *By using rational bases, our approach minimizes I by a two-step gradient descent which maintains the linearity of the whole formulation and drastically reduces the search effort.*

We first assume that all Z_i 's are equal, which is equivalent to approximating the *rational* bases using the corresponding *non-rational* bases. This allows us to estimate the 2D control vertex positions. Affine invariant parameters can be used as an initial estimate for point correspondence, which will be adjusted in succeeding steps to account for perspective foreshortening.

Observe that $dI = \sum_i (\frac{\partial I}{\partial \mathbf{p}'_i} d\mathbf{p}'_i + \frac{\partial I}{\partial R'_{i,k}} dR'_{i,k})$, suggesting that minimization can be broken into two stages: (1) that of updating 2D control vertex positions ($d\mathbf{p}'_i$); and (2) that of updating rational bases ($dR'_{i,k}$).

The estimated 2D control vertex positions are used to constrain the unknown rotation and translation parameters using Eq. 7. A *linear* formulation results using at least six 2D control vertices estimated from Eq. 6. (For a planar 3D curve, four 2D control vertex positions will suffice.) The motion parameters allow $R_{i,k}$'s to be updated using Eq. 8.

The updated $R_{i,k}$'s allow a better prediction of the appearance of the curve in images, and any discrepancy in the predicted and actual appearance of the curve is used in a gradient search to further verify the consistency. The prediction involves updating the parameterization t and the 2D control vertex positions \mathbf{p}_i , which are then used to estimate the unknown motion parameters through Eq. 7.

Hence, a recursive process results to refine the positions of the 2D control vertices, the shapes of the rational spline functions, the parameterization, and the 3D motion parameters, until a convergence is achieved.

Variation in Lighting Condition We now consider the case when the imaged objects are illuminated by light sources of different numbers, positions, and types. For simplicity, we will consider three spectral bands of red, green, and blue. Generalizing to an n -band illumination model is straightforward.

Assuming two 2D images differ only by scene illumination (*i.e.*, no geometrical changes), we can linearize *interesting (or important)* 2D regions by well-known techniques. We can then treat the problem as an illumination invariance problem for points along a curve.

In addition, we can include the affine or perspective case, to produce an invariant which is insensitive to both geometric (affine or perspective) and illumination changes. By solving for the deformation and translation parameters

from the affine or perspective invariants, we can reconstruct the same transformation for any point or curve between two images. Hence, any curve constructed from one image can be matched, point by point, to its corresponding curve in the transformed image. Illumination invariants for curves can then be applied, to verify if the two image regions, as the defined by the curves, are the same.

Let $\mathbf{L}(t)$ denote the perceived image color distribution along a curve. We have $\mathbf{L}(t) = [r(t), g(t), b(t)]^T = \int [f^r(\lambda), f^g(\lambda), f^b(\lambda)]^T s(\lambda, t) d\lambda$, where λ denotes the wavelength, and $f^r(\lambda)$ the sensitivity of the red sensor (similarly for the green and blue channels). We assume a Lambertian model, and that the reflected radiance functions, $s(\lambda, t)$, are modeled as a linear combination of a small number of basis functions $s_k(\lambda)$, whence, $s(\lambda, t) = \sum_k \alpha_k(t) s_k(\lambda)$, where $s_k(\lambda)$ denotes the k -th basis function for representing the reflected radiance properties, and $\alpha_k(t)$ are the space varying expansion coefficients. Then using an analysis which is similar to that employed in the affine case, we have

$$\mathbf{u}_{a,b} = \int \mathbf{L} \psi_{a,b} dt = \begin{bmatrix} L_1^r & L_2^r & \cdots & L_k^r \\ L_1^g & L_2^g & \cdots & L_k^g \\ L_1^b & L_2^b & \cdots & L_k^b \end{bmatrix} \begin{bmatrix} v_{a,b}^1 \\ \cdots \\ v_{a,b}^k \end{bmatrix} = \mathbf{L}_{rgb} \mathbf{v}_{a,b},$$

where $\begin{bmatrix} L_k^r \\ L_k^g \\ L_k^b \end{bmatrix} = \int_{\lambda} \begin{bmatrix} f^r(\lambda) s_k(\lambda) \\ f^g(\lambda) s_k(\lambda) \\ f^b(\lambda) s_k(\lambda) \end{bmatrix} d\lambda$ and $v_{a,b}^k = \int_t \alpha_k(t) \psi_{a,b} dt$.

Similarly,

$$\mathbf{u}'_{a,b} = \begin{bmatrix} L_1^{r'} & L_2^{r'} & \cdots & L_k^{r'} \\ L_1^{g'} & L_2^{g'} & \cdots & L_k^{g'} \\ L_1^{b'} & L_2^{b'} & \cdots & L_k^{b'} \end{bmatrix} \begin{bmatrix} v_{a,\pm b+t_0}^1 \\ \cdots \\ v_{a,\pm b+t_0}^k \end{bmatrix} = (\mathbf{L}'_{rgb})(\mathbf{v}_{a,\pm b+t_0}).$$

Then it is easily shown that the following expression is invariant under different lighting conditions (similar to Eq. 5):

$$\frac{\begin{bmatrix} \mathbf{u}'_{a_1, b_1} \cdots \mathbf{u}'_{a_k, b_k} \end{bmatrix}^T \begin{bmatrix} \mathbf{u}'_{a_1, b_1} \cdots \mathbf{u}'_{a_k, b_k} \end{bmatrix}}{\begin{bmatrix} \mathbf{u}'_{c_1, d_1} \cdots \mathbf{u}'_{c_k, d_k} \end{bmatrix}^T \begin{bmatrix} \mathbf{u}'_{c_1, d_1} \cdots \mathbf{u}'_{c_k, d_k} \end{bmatrix}} = \frac{\begin{bmatrix} \mathbf{u}_{a_1, \pm b_1+t_0} \cdots \mathbf{u}_{a_k, \pm b_k+t_0} \end{bmatrix}^T \begin{bmatrix} \mathbf{u}_{a_1, \pm b_1+t_0} \cdots \mathbf{u}_{a_k, \pm b_k+t_0} \end{bmatrix}}{\begin{bmatrix} \mathbf{u}_{c_1, \pm d_1+t_0} \cdots \mathbf{u}_{c_k, \pm d_k+t_0} \end{bmatrix}^T \begin{bmatrix} \mathbf{u}_{c_1, \pm d_1+t_0} \cdots \mathbf{u}_{c_k, \pm d_k+t_0} \end{bmatrix}} \quad (9)$$

3 Experimental Results

We conducted various experiments to test the validity of the new invariant features. Each experiment was isolated, which individually examined the performance of each image feature. However, the features can potentially be combined to make a powerful image retrieval system that can do *within-a-class* retrieval.

General Affine Transform with Change of Parameterization Fig. 1 shows (a) a shirt with a dolphin imprint and (b) a deformed version of the same imprint (an affine transformation). The extracted patterns are shown in (c). The second-order b-spline function of a uniform knot vector was used in the basis expansion. The invariant signatures shown in (d), which were aligned by maximizing the cross-correlation, are clearly quite consistent.

Perspective Transform Our formulation, though recursive in nature, is nonetheless linear and achieves fast convergence in our preliminary experiments. The number of iterations needed to verify the invariance was small (about 3 to 4) even for large perspective distortion. In Fig. 2, (a) shows the canonical view of a curve embedded on a curved surface (a cylindrical pail) and (b) another perspective. We extracted the silhouette of the car from both images and the depth values for the silhouette in canonic view were computed. Curve fitting and invariant signature (after five iterations) thus computed are displayed in Figs. 2(c) and (d), respectively. Our invariance framework produces consistent results for general, non-planar 3D curves, all with a small number of iterations.

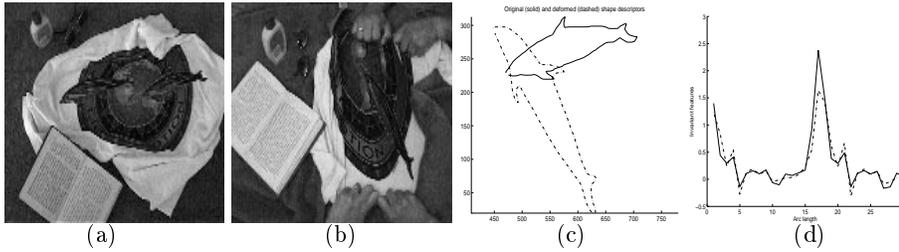


Fig. 1. (a) Original image, (b) deformed image, (c) extracted original (solid) and deformed (dashed) patterns, and (d) the invariant signatures plotted along the contours.

Change of Illumination To illustrate the correctness of the invariance formulation under illumination changes, we placed different color filters in front of the light sources used to illuminate the scene and verify the similarity of illumination invariant signatures. Fig. 3 shows the same cookbook cover under (a) white and (b) red illumination. For simplicity, we randomly defined two circular curves (indicated by the red and green circles) and computed the invariant signatures along these two curves under white and red illumination. It should be noted that the particular example we show here only serve to demonstrate the correctness of the framework. In real applications, we can linearize the image to obtain an invariant signature for the *whole* image. The invariant profiles computed from the white (solid) and red (dashed) illumination are shown in Fig. 3(c) for the curve defined by the red circle and (d) for the curve defined by the green circle. As can be seen from the figure, the signatures are quite consistent.

Hierarchical Invariant Analysis The additional degree of freedom in designing the basis function enables a hierarchical shape analysis. Fig. 4(a) shows the original and noise-corrupted shapes. As shown in Fig. 4(b)-(c), our approach, which analyzes the shape at different scales locally, will eventually discover the similarity, even though the similarity may manifest at different levels of details. In this case, scale 8 produces more consistent signatures than the others.

Future Work The performance of each image feature is very encouraging, prompting us to combine these image features to make a powerful image retrieval system that can do *within-a-class* retrieval. Results will be presented in a future paper. Applications include searching through specialized image databases, which contains imaged objects with very similar appearance (*e.g.*, botanical databases and aquarium databases). In fact, these features have already been applied to object

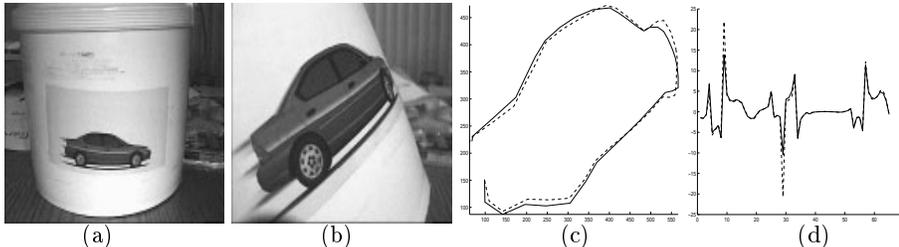


Fig. 2. (a) Canonical view, (b) another perspective, (c) 2D image curve (solid) and the curve derived w. perspective invariant fitting (dashed), and (d) their shape signatures.

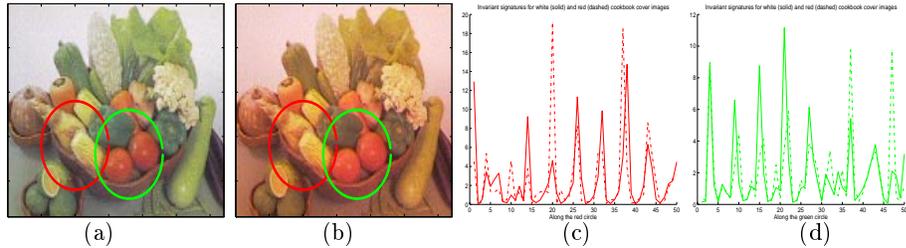


Fig. 3. The same cookbook cover under (a) white and (b) red illumination, and the invariant signatures computed under white (solid) and red (dashed) illumination (c) along the red circle and (d) along the green circle.

recognition experiments where perspective distortion, color variation, noise, and occlusion were all present [1]. In that experiment, the database comprised of different models of airplanes, many of which had the same general shape. Perfect recognition was achieved for that particular database and test images.

4 Conclusion

We presented a new framework for computing image-derived, invariant features, ideal for image indexing and retrieval. These features provide high discriminative power and are insensitive to many environmental changes. Preliminary results show promise as a useful tool for searching image databases.

References

1. R. Alferéz and Y.F. Wang. Geometric and Illumination Invariants for Object Recognition. *IEEE Trans. Pattern Analy. Machine Intell.* To appear as a regular paper.
2. K. Arbter, W. E. Snyder, H. Burkhardt, and G. Hirzinger. Application of Affine-Invariant Fourier Descriptors to Recognition of 3-D Objects. *IEEE Trans. Pattern Analy. Machine Intell.*, 12:640–647, 1990.
3. I. Daubechies. Orthonormal Bases of Compactly Supported Wavelets. *Commun. Pure Appl. Math.*, 41:909–960, 1988.
4. Hampapur et.al. Virage Video Engine. *Proc. of SPIE, Storage and Retrieval for Image and Video Databases V*, 3022:188–200, 1997.
5. M. Flickner et.al. Query by Image and Video Content: The QBIC System. *IEEE Comput.*, pages 23–32, September 1995.
6. A. Pentland, R.W. Picard, and S. Sclaroff. Photobook: Tools for Content-Base Manipulation of Image Databases. *Int. J. Comput. Vision*, 18(3):233–254, 1996.
7. T. H. Reiss. *Recognizing Planar Objects Using Invariant Image Features*. Springer-Verlag, Berlin, 1993.
8. I. Weiss. Geometric Invariants and Object Recognition. *Int. J. Comput. Vision*, 10(3):207–231, 1993.

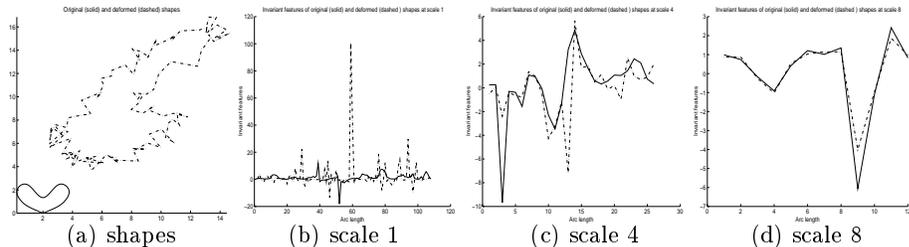


Fig. 4. Invariant shape descriptors for the original (solid) and deformed, noise-corrupted shapes (dashed) at different scales.