

CLASSIFICATION OF UNLABELED POINT SETS USING ANSIG

José J. Rodrigues, João M. F. Xavier, Pedro M. Q. Aguiar

Institute for Systems and Robotics / IST, Lisboa, Portugal
jerasman@gmail.com, {jxavier, aguiar}@isr.ist.utl.pt

ABSTRACT

We address two-dimensional shape-based classification, considering shapes described by arbitrary sets of unlabeled points, or landmarks. This is relevant in practice because, in many applications, the points describing the shapes come from automatic processes, *e.g.*, edge detection, thus without labels. Rather than attempting to compute point correspondences (a quagmire, when dealing with nontrivial shapes), we use what we call the analytic signature (ANSIG) of the shapes, a representation that has the key feature of being invariant to point labeling. Geometric transformations, such as translation, rotation, and scale, and different cardinality of point sets, are also dealt with by this representation. We demonstrate the capabilities of our representation with several shape classification experiments.

Index Terms— Shape classification, shape representation, unlabeled data, permutation invariance, analytic signature, ANSIG.

1. INTRODUCTION

Many objects are primarily recognized by their shape, rather than their color or texture. Although this fact has been confirmed by surveys showing that users would prefer to retrieve images from shape queries [1], the majority of content-based image retrieval systems still use color and texture features to compare images. In fact, shape-based classification proved to be a very hard task, remaining an open problem, which we address in this paper.

We consider two-dimensional (2D) shapes described by arbitrary sets of *unlabeled* points, or landmarks. The challenge in shape classification is how to represent these sets. Fig. 1 illustrates the great diversity that can arise when dealing with arbitrary shapes. While the leftmost shape seems adequate to be described by its contour, for the middle one, the outside contour would not certainly be enough, and the rightmost shape requires a more general representation scheme. This motivates our work with generic sets of points.

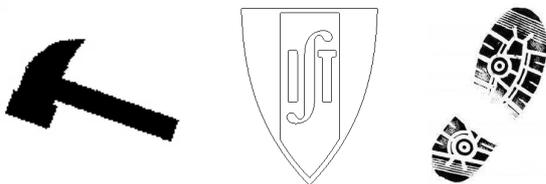


Fig. 1. 2D shape diversity. Left: a “blob”; middle: shape characterized by several contours; right: a more general set of points.

Partially supported by FCT, under ISR/IST plurianual funding (POSC program, FEDER), and grants MODI-PTDC/EEA-ACR/72201/2006 and SIPM-PTDC/EEA-ACR/73749/2006.

The *statistical theory of shape* [2], developed for shapes described by sets of *labeled* landmarks, copes with geometric transformations and shape variations. Although it has led to significant results, in many practical scenarios the correspondences between shape landmarks are not available. The majority of the methods to deal with unlabeled point sets focus on representing a “blob”, *i.e.*, a connected set of points, like the leftmost example in Fig 1. Among these, *region-based* methods use moment descriptors, *e.g.*, geometrical, Legendre, Zernike [3], or Tchebichef [4], and *contour-based* methods represent the boundary of the shape using, *e.g.*, curvature scale space [5], wavelets [6], contour displacements [7], splines [8], or Fourier descriptors [9]. Although (some of) these representations are invariant to geometric transformations, they are restricted to shapes well described by closed contours. Since shape cues come often from image edges and it is hard to extract complete contours when dealing with real images, researchers also developed local shape descriptors that, at each point of the shape, capture the relative distribution of the remaining points, *e.g.*, *shape contexts* [10] and *distance multisets* [11]. These representations do not deal with general shapes and geometric transformations. A number of approaches to deal with general sets of unlabeled points are motivated by the need to compute the rigid transformation between them. To cope with unlabeled points, they use iterative algorithms that compute, in alternate steps, the registration parameters and the point correspondences. One of the better known examples is the *Iterative Closest Point (ICP)* algorithm [12]. More recently, researchers have proposed statistical methods that use “soft” correspondences [13, 14], leading to *Expectation-Maximization (EM)*-like iterative algorithms that, although having been used in challenging scenarios, *e.g.*, part decomposition and nonrigid registration, have the common limitations of uncertain convergence and sensitivity to initialization.

In shape-based classification, rather than attempting to compute the correspondences between the points describing the shapes, we seek to represent shape in a way that is invariant to the point labels (besides relevant geometric transformations, like translation, rotation, and scale). To achieve this, we use the representation introduced in [15]: the analytic signature (ANSIG). Here, a shape is represented by an analytic function on the complex plane. Shapes that differ by a re-ordering of the set of landmarks have the same ANSIG, *i.e.*, this representation is *permutation invariant*. Furthermore, this representation is discriminative, *i.e.*, different shapes are represented by different ANSIGs. Under our approach, shape-based classification boils down to comparing the ANSIGs of candidate shapes with the ones in a shape database. As any analytic function, the ANSIG is completely described by the values it takes in a closed contour on the complex plane. Thus, we store an ANSIG by sampling it in the unit circle.

Permutation invariance has been recently addressed: in [16, 17], permutation is factored out after being computed as the solution of a convex optimization problem, an approach that does not deal with

geometric transformations, such as rotation; in [18], a shape is described by the set of distances between all pairs of points, limiting the applicability to small sets of landmarks, like fingerprint minutiae.

2. THE ANALYTIC SIGNATURE OF A SHAPE

Consider a 2D shape described by a set of n unlabeled points, or landmarks, in the plane \mathbb{R}^2 . Under the usual identification of \mathbb{R}^2 with the complex plane \mathbb{C} , that shape can be represented by the vector

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} x_1 + iy_1 \\ x_2 + iy_2 \\ \vdots \\ x_n + iy_n \end{bmatrix} \in \mathbb{C}^n,$$

or by any other vector related to \mathbf{z} by a permutation of its entries. To cope with this quagmire of different orderings, we propose to represent the shape \mathbf{z} by what we call its ANSIG, an analytic function $a(\mathbf{z}, \cdot)$, defined by

$$a(\mathbf{z}, \xi) = \frac{1}{n} \sum_{m=1}^n e^{z_m \xi}, \quad (1)$$

where ξ is a dummy complex variable. We will see that the ANSIG (1) exhibits properties that make it adequate for shape representation, in particular, it is invariant with respect to the labeling of the landmarks representing the shape.

It is clear that the ANSIG is invariant to the labeling of the landmarks because, from the definition of a in (1), it follows that $a(\mathbf{z}, \cdot) = a(\mathbf{w}, \cdot)$ whenever \mathbf{z}, \mathbf{w} are related by a permutation. Furthermore, the ANSIG is a maximal invariant, in the sense that if two shapes have equal ANSIGs, they must be the same. To prove this, consider two equal ANSIGs, $a(\mathbf{z}, \cdot) = a(\mathbf{w}, \cdot)$. Naturally, their derivatives also coincide. In particular, when evaluated at the origin:

$$\left. \frac{d^k}{d\xi^k} a(\mathbf{z}, \xi) \right|_{\xi=0} = \left. \frac{d^k}{d\xi^k} a(\mathbf{w}, \xi) \right|_{\xi=0} \quad k = 1, 2, \dots, n. \quad (2)$$

Using the definition of $a(\cdot, \cdot)$ in (1), the system of equations (2) is written in terms of the entries of vectors \mathbf{z} and \mathbf{w} as

$$\begin{cases} z_1 + z_2 + \dots + z_n = w_1 + w_2 + \dots + w_n \\ z_1^2 + z_2^2 + \dots + z_n^2 = w_1^2 + w_2^2 + \dots + w_n^2 \\ \vdots \\ z_1^n + z_2^n + \dots + z_n^n = w_1^n + w_2^n + \dots + w_n^n. \end{cases}$$

This set of equalities implies that the polynomials $p(t) = (t - z_1)(t - z_2) \dots (t - z_n)$ and $q(t) = (t - w_1)(t - w_2) \dots (t - w_n)$ are identical, see [19], sharing the same system of roots (including multiplicities), thus the (multi-)sets $\{z_1, z_2, \dots, z_n\}$ and $\{w_1, w_2, \dots, w_n\}$ are equal. This way we conclude that the vectors $\mathbf{z} = [z_1 z_2 \dots z_n]^T$ and $\mathbf{w} = [w_1 w_2 \dots w_n]^T$ are equal up to a permutation, proving the maximal invariance of our ANSIG.

We now illustrate this characteristic of the ANSIG representation using two shapes whose only difference is the order by which the landmarks are stored. In the top row of Fig. 2, we represent such a pair of shapes. Note that the set of landmarks (black dots) for the left image is the same than that of the right one. To indicate the order by which the landmarks are stored, we use a line connecting them, thus the images look different, illustrating the quagmire of having to compare shapes when the correspondence between the sets of landmarks is unknown. We then compute the ANSIG of each

of these two shapes. Although the vectors containing the landmarks for these shapes are completely different (due to re-ordering), their ANSIG results equal, as expected from our derivations above, and illustrated by the plot in the middle of Fig. 2, which shows the common ANSIG¹. We must also emphasize that, as demonstrated above, the ANSIG of geometrically different shapes results in fact distinct.

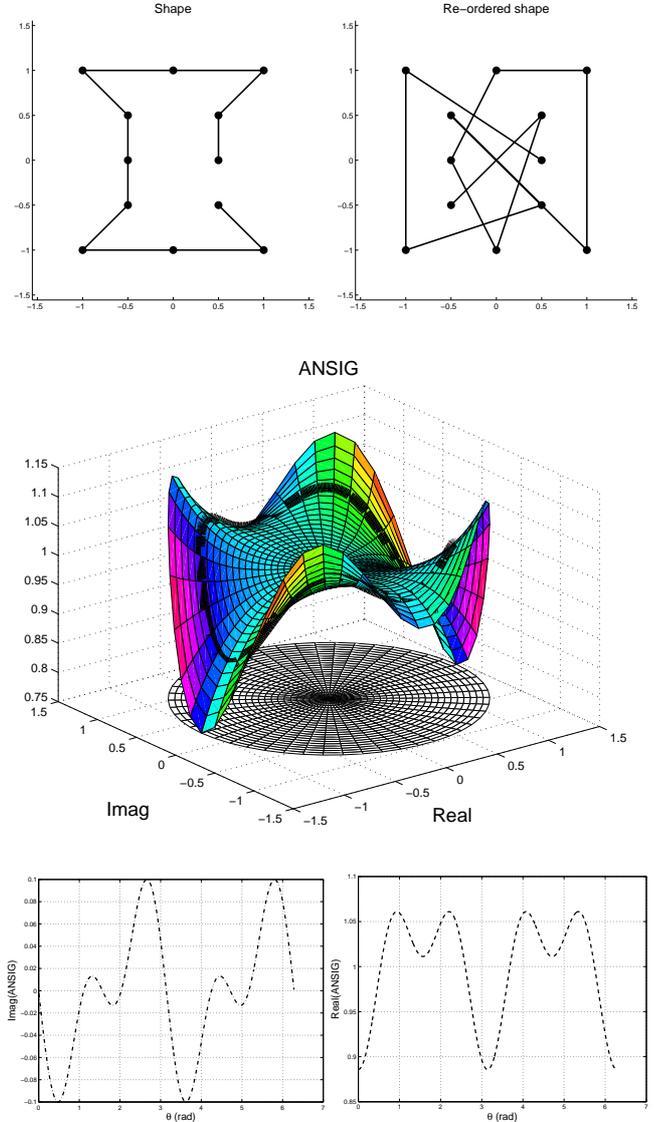


Fig. 2. Permutation invariance of the ANSIG representation. Top: two representations of the same shape, described by the positions of a common set of 12 points. The line connecting the dots indicates the order by which the points are stored. Middle: ANSIG of any of the representations above. Bottom: ANSIG characterization by means of the real and imaginary parts of its restriction to the unit-circle.

The ANSIG in (1) maps a shape, *i.e.*, a vector \mathbf{z} to an analytic function $a(\mathbf{z}, \cdot)$ on the complex plane, what could seem inadequate to a computer implementation. However, a major consequence of Cauchy's integral formula is that any analytic function

¹The plot uses a common representation for complex functions, where the real part is represented by surface height and the imaginary part by its color.

is unambiguously determined by the values it takes on a simple closed contour, see, *e.g.*, [20]. In particular, if we choose the contour to be the unit-circle (the circle of radius 1 centered at the origin), an analytic function f is uniquely determined by $\{f(e^{j\varphi}) : \varphi \in [0, 2\pi]\}$. In practice, this is approximated by sampling, *i.e.*, by storing the values of f on a finite set of points in the unit-circle, $\{1, W_K, W_K^2, \dots, W_K^{K-1}\}$, where

$$W_K = e^{j\frac{2\pi}{K}}. \quad (3)$$

The integer K in (3) controls the sampling density². In summary, we approximate the ANSIG $a(\mathbf{z}, \cdot)$ by its discrete counterpart

$$\mathbf{a}_K(\mathbf{z}) = [a(\mathbf{z}, 1) \quad a(\mathbf{z}, W_K) \quad \dots \quad a(\mathbf{z}, W_K^{K-1})]^T. \quad (4)$$

The process to store an ANSIG is also illustrated in Fig. 2: the unit-circle sampling is represented by the thick black line in the middle plot and the corresponding discretized ANSIG \mathbf{a}_K is represented in the bottom plots (real and imaginary parts). Although this process resembles the one of obtaining the discrete-time Fourier Transform from the Z-transform, our ANSIG has nothing to do with Fourier-related representations, such as the ones using Fourier Descriptors [9]. In fact, while Fourier analysis is concerned with how patterns evolve across time (across landmark index, in shape description), our ANSIG is not – it is even invariant to the order the landmarks are stored. We postpone a comment on the fact the ANSIG plots in the bottom line of Fig. 2 are periodic, to after describing the effect of geometric transformations.

3. SHAPE-BASED CLASSIFICATION

In shape-based classification, besides the key issue of invariance to point labeling, it is also necessary to deal with geometric transformations like translation, rotation, and scale. To factor out translation and scale, we pre-process the shape vector \mathbf{z} , by centering and normalizing it, working then with the ANSIG $a(\tilde{\mathbf{z}}, \cdot)$, being

$$\tilde{\mathbf{z}} = \sqrt{n} \frac{\mathbf{z} - \bar{\mathbf{z}}}{\|\mathbf{z} - \bar{\mathbf{z}}\|}, \quad (5)$$

where $\bar{\mathbf{z}}$ is a vector with all entries equal to the mean of \mathbf{z} , and $\|\cdot\|$ denotes the 2-norm. It is obvious that two shapes \mathbf{z} and \mathbf{w} that differ by a translation and a scale factor, lead to $\tilde{\mathbf{z}} = \tilde{\mathbf{w}}$, as desired.

The reader may wonder why the factor \sqrt{n} in (5). In fact, this factor has nothing to do with factoring out translation or scale – it is constant for shapes described by the same number n of landmarks. The motivation for that factor is precisely to enable dealing well with shapes described by different numbers of landmarks, providing robustness to over/under sampling the shapes. To illustrate this capability, consider an example where each landmark in the n -dimensional vector \mathbf{z} is repeated p times and stored in the pn -dimensional vector \mathbf{z}_p , *i.e.*, $\mathbf{z}_p = \mathbf{1}_p \otimes \mathbf{z}$, where \otimes is the Kronecker product. It is clear that vectors \mathbf{z} and \mathbf{z}_p represent the same shape. This example, although extreme, illustrates what may happen when comparing shapes that, although similar, have been differently sampled, leading to vectors of different dimensions, what may happen, for example, when the shapes to compare come from edge maps (automatically) detected from images of different resolutions. We now show that the normalization factor \sqrt{n} in (5) guarantees that the same ANSIG will be produced. In fact, we have

$$\mathbf{z}_p - \bar{\mathbf{z}}_p = \mathbf{1}_p \otimes (\mathbf{z} - \bar{\mathbf{z}}) \quad \text{and} \quad \|\mathbf{z}_p - \bar{\mathbf{z}}_p\| = \sqrt{p} \|\mathbf{z} - \bar{\mathbf{z}}\|,$$

²In all the experiments reported in this paper, we used $K = 512$.

thus, the normalized version of \mathbf{z}_p is

$$\tilde{\mathbf{z}}_p = \sqrt{pn} \frac{\mathbf{z}_p - \bar{\mathbf{z}}_p}{\|\mathbf{z}_p - \bar{\mathbf{z}}_p\|} = \sqrt{n} \frac{\mathbf{1}_p \otimes (\mathbf{z} - \bar{\mathbf{z}})}{\|\mathbf{z} - \bar{\mathbf{z}}\|} = \mathbf{1}_p \otimes \tilde{\mathbf{z}}$$

and its ANSIG $a(\tilde{\mathbf{z}}_p, \cdot)$ can be successively written as

$$\begin{aligned} a(\tilde{\mathbf{z}}_p, \xi) &= a(\mathbf{1}_p \otimes \tilde{\mathbf{z}}, \xi) \\ &= \frac{1}{pn} \underbrace{\left(\sum_{m=1}^n e^{\tilde{\mathbf{z}}_m \xi} + \dots + \sum_{m=1}^n e^{\tilde{\mathbf{z}}_m \xi} \right)}_{p \text{ times}} \\ &= \frac{1}{n} \sum_{m=1}^n e^{\tilde{\mathbf{z}}_m \xi} \\ &= a(\tilde{\mathbf{z}}, \xi), \end{aligned} \quad (6)$$

establishing the desired equality between the signatures of the differently sampled shapes in \mathbf{z} and \mathbf{z}_p . In (6) and (7) we used (1).

We now look at how the rotation of a shape affects its ANSIG. Consider a shape \mathbf{z} rotated by an angle θ , *i.e.*, the shape $e^{j\theta} \mathbf{z}$. From definition (1), it follows that the ANSIG of the rotated shape is

$$a(e^{j\theta} \mathbf{z}, \xi) = \frac{1}{n} \sum_{k=1}^n e^{e^{j\theta} z_k \xi} = \frac{1}{n} \sum_{k=1}^n e^{z_k (e^{j\theta} \xi)} = a(\mathbf{z}, e^{j\theta} \xi).$$

Thus, the ANSIG of the rotated shape is simply a rotated version (in the complex plane) of the original one. Obviously, the unit-circle restriction of the rotated shape's ANSIG is then a simple circular shift of the original one. In particular, if the rotation matches one of the sample points in the uniform grid, *i.e.*, if $e^{j\theta} = W_K^k$ for some $k = 0, 1, \dots, K-1$, then the ANSIG vectors (4) are related by

$$\mathbf{a}_K(e^{j\theta} \mathbf{z}) = \mathbf{a}_K(\mathbf{z}) \bmod k, \quad (8)$$

where the notation $\bmod k$ stands for a k -step cyclic shift.

We now comment on the fact that the ANSIG plots in the bottom of Fig. 2 are periodic. This is due to the rotational symmetry of the shapes in the top. In fact, that particular shape exhibits mirror symmetry, being then invariant to rotations of multiples of π . As just seen above, shape rotation leads to a (circular) translation of the ANSIG plots. Thus, for this shape, the ANSIG plots are invariant to translations of multiples of π , *i.e.*, they are periodic with period π . This is confirmed by the plots in the bottom of Fig. 2 (and also by the π -rotational symmetry of the ANSIG representation in the middle).

Naturally, with K sufficiently large, any rotation is well approximated by a point in the sampling grid and (8) holds for practical purposes. Thus, to use ANSIG in shape-based classification, when comparing shapes \mathbf{z} and \mathbf{w} , we just have to compute the circular shift k^* that best “aligns” the ANSIG vectors of these two shapes, *i.e.*,

$$k^* = \arg \min_{k=0,1,\dots,K-1} \|\mathbf{a}_K(\mathbf{z}) - \mathbf{a}_K(\mathbf{w}) \bmod k\|^2, \quad (9)$$

and then evaluate the similarity between $\mathbf{a}_K(\mathbf{z})$ and $\mathbf{a}_K(\mathbf{w}) \bmod k^*$. Our experiments have shown that a good choice for measure of similarity is the (cosine of) the angle between the ANSIG vectors, *i.e.*,

$$\text{similarity}(\mathbf{z}, \mathbf{w}) = \frac{|\mathbf{a}_K^H(\mathbf{z}) \mathbf{a}_K(\mathbf{w}) \bmod k^*|}{\|\mathbf{a}_K(\mathbf{z})\| \|\mathbf{a}_K(\mathbf{w})\|},$$

where \mathbf{a}_K^H denotes the conjugate transpose of \mathbf{a}_K .

Finally, we remark that, although solving (9) by exhaustive search leads to an algorithm of computational complexity $\mathcal{O}(K^2)$ (which is incomparably smaller to the complexity of considering the $n!$ possible landmark re-orderings), it is straightforward to reduce its complexity to $\mathcal{O}(K \log K)$ by using the Fast Fourier Transform.

4. EXPERIMENTS

To illustrate the versatility of our representation, we use a database of 56 “clip-art” binary images, which contains several complex shapes, see Fig. 3. Our experiments consisted in attempting to match noisy observations of (randomly) translated, rotated, and scaled versions of these shapes, to the correct entry in the database, using the simple scheme outlined in the previous section, *i.e.*, just selecting the closest ANSIG. We performed 200 tests per shape and for each of the two noise levels illustrated in Fig. 4. Although some of the test shapes were so distorted that would be hardly recognizable by humans, we obtained 98.4% correct classification for the higher level of noise and 100% for the other (the only systematic error was confusing a pen with the baseball bat, which have in fact very similar shapes as shown in the top right corner of Fig. 3). We performed similar tests with the MPEG-7 shape database, obtaining similar results.

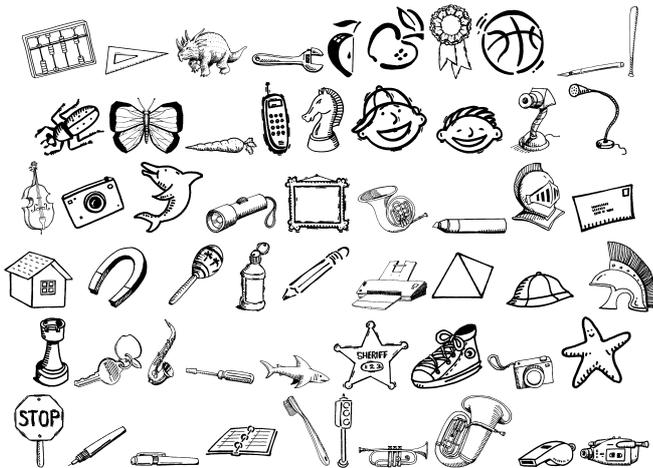


Fig. 3. Clip-art database of binary images.

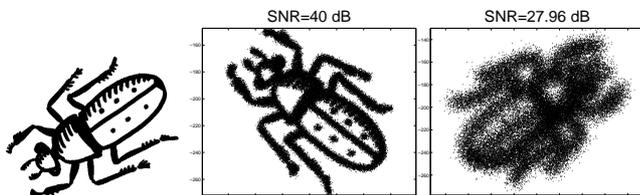


Fig. 4. Examples of noisy test shapes used for classification (middle and right) and corresponding noiseless image in the database (left).

In a different scenario, we use a database of 20 shapes of real objects, obtained by performing standard edge detection on real images, see Fig. 5. We have classified a distinct set of approx. 100 images, obtained with an off-the-shelf webcam, see examples in Fig. 6, by comparing the ANSIGs of their edge maps with those of the shapes in Fig. 5. In spite of incomplete edges and diverse sizes, positions, and orientations, all images were correctly classified. Other successful experiments included shape-based trademark retrieval.

5. CONCLUSION

We proposed a simple approach to the classification of 2D shapes. The distinctive feature of our method is its ability to deal with shapes described by arbitrary sets of unlabeled points. Representing these sets by what we call their analytic signatures (ANSIGs), shape classification is done without any search for point correspondences.

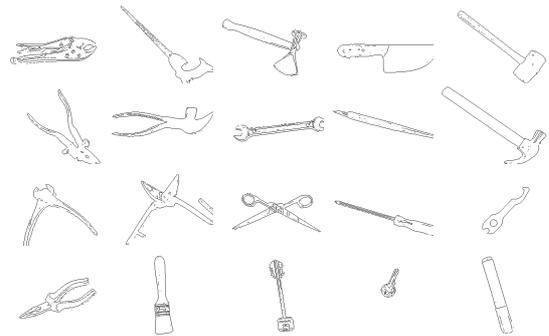


Fig. 5. Database of shapes of real objects.

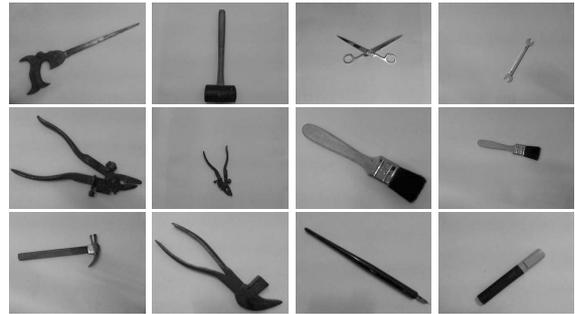


Fig. 6. Examples of webcam images used for classification.

6. REFERENCES

- [1] L. Schomaker, E. Leau, and L. Vuurpijl, “Using pen-based outlines for object-based annotation and image-based queries,” in *Int. Conf. Visual Information and Information Systems*, Amsterdam, 1999.
- [2] D. Kendall et al, *Shape and Shape Theory*, John Wiley and Sons, 1999.
- [3] A. Khotanzan and Y. Hong, “Invariant image recognition by Zernike moments,” *IEEE T-PAMI*, 12:489–497, 1990.
- [4] R. Mukundan, S. Ong, and P Lee, “Image analysis by Tchebichef moments,” *IEEE T-IP*, 10:1357–1364, 2001.
- [5] F. Mokhtarian and A. Mackworth, “A theory of multiscale, curvature-based shape representation for planar curves,” *IEEE T-PAMI*, 14, 1992.
- [6] G. Chuang and C. Kuo, “Wavelet descriptor of planar curves: Theory and applications,” *IEEE T-IP*, 5:56–70, 1996.
- [7] T. Adamek and N. O’Connor, “A multiscale representation method for nonrigid shapes with a single closed contour,” *IEEE T-PAMI*, 14, 2004.
- [8] Paul Dierckx, *Curve and Surface Fitting with Splines*, Oxford Science Publications, 1995.
- [9] I. Bartolini, P. Ciaccia, and M. Patella, “Warp: Accurate retrieval of shapes using phase of fourier descriptors and time warping distance,” *IEEE T-PAMI*, 27(1):142–147, 2005.
- [10] S. Belongie, J. Malik, and J. Puzicha, “Shape matching and object recognition using shape contexts,” *IEEE T-PAMI*, 24(24), 2002.
- [11] C. Grigorescu and N. Petkov, “Distance sets for shape filters and shape recognition,” *IEEE T-IP*, 12(10):1274–1286, 2003.
- [12] P. Besl and N. McKay, “A method for registration of 3-D shapes,” *IEEE T-PAMI*, 14(2), 1992.
- [13] H. Chui and A. Rangarajan, “A new algorithm for non-rigid point matching,” in *IEEE CVPR*, Hilton Head Island SC, USA, 2000.
- [14] G. McNeill and S. Vijayakumar, “A probabilistic approach to robust shape matching,” in *IEEE ICIP*, Atlanta GA, USA, 2006.
- [15] J. Rodrigues and P. Aguiar and J. Xavier, “ANSIG – an analytical signature for permutation-invariant two-dimensional shape representation,” in *IEEE CVPR*, Anchorage, Alaska, USA, 2008.
- [16] T. Jebara, “Images as bags of pixels,” in *IEEE CVPR*, Nice, 2003.
- [17] R. Kondor, A. Howard, and T. Jebara, “Multi-object tracking with representations of the symmetric group,” in *Int. Conf. on Artificial Intelligence and Statistics*, San Juan, Puerto Rico, 2007.
- [18] M. Boutin and M. Comer, “Faithful shape representation for 2D gaussian mixtures,” in *IEEE ICIP*, San Antonio TX, USA, 2007.
- [19] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, UK, 1985.
- [20] L. Ahlfors, *Complex Analysis*, McGraw-Hill, USA, 1978.