

A BLIND DEFINITION OF SHAPE

J.L. LISANI¹, J.M. MOREL² AND L. RUDIN³

Abstract. In this note, we propose a general definition of shape which is both compatible with the one proposed in phenomenology (gestaltism) and with a computer vision implementation. We reverse the usual order in Computer Vision. We do not define “shape recognition” as a task which requires a “model” pattern which is searched in all images of a certain kind. We give instead a “blind” definition of shapes relying only on invariance and repetition arguments. Given a set of images \mathcal{I} , we call shape of this set any spatial pattern which can be found at several locations of some image, or in several different images of \mathcal{I} . (This means that the shapes of a set of images are defined without any *a priori* assumption or knowledge.) The definition is powerful when it is invariant and we prove that the following invariance requirements can be matched in theory and in practice: local contrast invariance, robustness to blur, noise and sampling, affine deformations. We display experiments with single images and image pairs. In each case, we display the detected shapes. Surprisingly enough, but in accordance with Gestalt theory, the repetition of shapes is so frequent in human environment, that many shapes can even be learned from single images.

Mathematics Subject Classification. 62H35.

Received January 13, 2002.

1. INTRODUCTION

1.1. Phenomenology of shape

The problem of how humans and animals identify visually shapes has been studied at the beginning of the past century by phenomenologists, ethologists and psychologists. The mainstream of this research is summarized in the Bible of the Gestalt theory [24] and in Kanizsa [16]. In these books, not only the capacity of humans to recognize instantaneously similar shapes in an image is stressed, but this capacity is ranked as the main *grouping* principle, *i.e.* the main phenomenal process by which global entities can be seen in an image.

We have here two complementary mechanisms: first, similar shapes in an image are detected and, second, they get *ipso facto* grouped into a higher level entity, a *Gestalt*. Thus, the geometric ability of humans (and most mammals) to compare automatically and in very short time shapes appearing in an image is supported by phenomenology. The main point which was stressed is this: this recognition process is widely independent of any *a priori* knowledge of the seen shapes. This is contrary to Platon’s conjecture on the mechanism of recognition as the matching of a perceived object to an “idea” of the object. We have a geometric ability to match shapes which we have never seen before. This ability is, however, hindered when shapes are presented

Keywords and phrases: Image analysis, basic shape elements, contrast invariance, level lines, scale space.

¹ Univ. de les Illes Balears, Cra. de Valldemossa, km 7.5, 07071 Palma, Spain; e-mail: vdmi1r0@clust.uib.es

² CMLA, ENS Cachan, 61 avenue du Président Wilson, 94235 Cachan, France; e-mail: more1@cmla.ens-cachan.fr

³ Cognitech Inc., 225 S. Lake Avenue, CA-91101 Pasadena, U.S.A.; e-mail: lenny@cognitech.com

out of context. This fact is used, as a game, in the Images d'Epinal containing hidden shapes: the familiar silhouette of Napoleon is hard to detect when it is displayed as a free space between the branches of a tree.

The human shape recognition performance also decreases when the shapes change strongly orientation or scale. It is, however, robust under such drastic perturbations as Gaussian noise, contrast or color changes, and affine or projective deformations.

The recognition of an "object" in a new image is more of a psychological and philosophical nature: how do we "recognize" a chair whose model we have never seen before? Clearly, we have an "idea" or a concept of what a chair is and the decision "this is a chair" highly depends upon a wide context, including considerations like position of the object in the scene, and involving our common sense. For the time being, this kind of recognition process is out of range for mathematical modelling.

Quite different and more treatable looks the following problem: a photograph contains two or more chairs of exactly the same design and wish to detect this fact automatically. Of course, this detection would be in no way a "recognition of a chair". Now, it is a preliminary important step to the general recognition problem: according to gestaltists, we are able to recognize identical objects, no matter what their meaning and shape is.

1.2. Shape from the Computer Vision viewpoint

Computer Vision, a discipline born in the seventies, has addressed the problem of performing "visual tasks" by computer algorithms applied to digital images. The shape recognition problem, understood as the comparison of two identical objects captured in different images, has received two different kinds of answers.

A first answer, industry oriented, has been to give up any general theory of shape, but to focus on shapes of particular interest like fingerprints, mechanical parts of machines, printed characters, handwritings of addresses on envelopes or signatures on checks, etc. For such objects, specific techniques, relying on human expertise and a lot of statistical experience, have led to significant successes: software by now exists for efficient fingerprint identification and for printed text.

A second, more ambitious direction is to build a general theory of shape recognition. Here, the state of the art is by far more blurry. To the best of our knowledge, there is no software (and no theory) addressing the automatic recognition of any 3D object. There has been for a long time a relatively complete theory only in one case: the recognition of planar shapes seen "without occlusion". This means that the shape to be recognized must be completely under view (no hidden or missing parts).

Let us mention the existence of several techniques for the description and comparison of global planar shapes in an image. All compute variants of moments or Fourier coefficients. Some techniques are based on quantitative measurements of the characteristics of the shapes [11, 31], such as perimeter, algebraic moments [15, 27, 35], Fourier coefficients [17, 18, 30], etc. Most of these techniques just permit a global registration of the shapes and the extracted features are sensitive to noise. Another related technique is normalization [31]. Normalization methods allow the transformation of any element of an equivalence class of shapes (under the action of some geometric group) into a specific one, fixed once for ever in each class. One can find a good account of affine invariant shape comparison techniques in [8].

These methods assume shapes to be wholesome. Now, the occlusion phenomenon is as universal in vision as the principle of linear superposition of sounds in audio: most objects we "see" are partially hidden by other objects which are in front, and each object we "see" also is partially occulting several objects in back. This problem is called "occlusion" in Computer Vision and known as the *figure-background* problem in phenomenology.

A non occlusion assumption is a severe restriction, which is not even acceptable in the most favorable case of character recognition: even there, experimental evidence shows that characters are in no way easy to single out. They glue together or fall apart in a very uncontrolled manner. Thus, shape recognition and comparison must be local and partial.

1.3. The shape extraction/recognition dilemma

While shape extraction looks trivial (but is not) in such cases as fingerprints or characters, it is highly nontrivial, and unsolved, in all usual photographs and digital images. Should we first "extract" the shapes and

then “recognize” them, or is recognition prior to, or simultaneous with extraction? This dilemma is usually solved by performing an “edge detector” in the image which extracts the points with high gradient. These points can be linked in many different ways to build shapes... but: a combinatorial blow up immediately ensues.

2. A BLIND DEFINITION OF SHAPE ELEMENTS

In this section, we first address the problem of deciding which information in a digital image is relevant to define or recognize shapes. This will be a preliminary step towards the definition of shape we shall give in continuation. We shall show that invariance arguments enforce the definition of the most basic shape elements.

2.1. The local contrast invariance argument

We define a digital image as a function $u(x)$, where $u(x)$ represents the grey level or luminance at x . (Shape recognition algorithms do not use much color. The information brought by color usually is redundant with respect to luminance [7].) Shape recognition algorithms are widely independent of global contrast changes of the image. Such contrast changes can be modelled as $g(u(x))$ where g is some increasing function. Thus shape must be encoded as information independent of g . This led the Mathematical Morphology school [34] to split the image into its level sets $\chi_\lambda u = \{x, u(x) \geq \lambda\}$. When we change the contrast of an image, its set of level sets remains the same. A further step was taken in [6]. In this paper, the authors proposed to directly use the set of level lines of the image. Assuming u is smooth enough, *e.g.* C^1 , the encoding of shapes by level lines permits to assume a stronger invariance criterion. Once singular levels are eliminated, the level lines are all closed Jordan curves which do not meet. Each one of these Jordan curves is a better candidate as shape element. The mentioned authors define and give a characterization of the local contrast changes as maps $x \rightarrow g(x, u(x))$ which maintain the same set of level lines (the same topographic map) for u and $g(x, u)$. We must retain from their construction the intuitive fact that local contrast changes reflect well enough the physical fact that objects in an image may change independently their grey level or color, without affecting their recognition. Thus, the application $u \rightarrow \text{set of level lines of } u$ is a good candidate as a shape parser for the image.

The role of the contrast invariance requirement has never been fully acknowledged in Computer Vision. As a consequence, most (and all, as far as we know) image comparison techniques described in the literature [5] fail to find common shapes in two different images when the illumination conditions are different. This is the case for the two main groups of image comparison methods which we now list, correlation and edge detection.

- **Correlation techniques.** They are based on the comparison, by a quadratic distance of grey levels [32], of rectangular windows of fixed or adaptive size [29], for the two images. Since the measure of the difference between the windows is based on the comparison of their grey levels, and the scale of grey levels can vary a lot from image to image, they cannot be used for shape comparison. They work decently with image sequences, where each frame is very similar to the previous and next ones.
- **Edge based methods.** Edges are arrays of points where the gradient is highest, as defined by Marr [22]. The edge detection algorithms depend upon a threshold that is hard to estimate, and usually fixed arbitrarily. The result of an edge detector is a set of points in the image, not organized in any structure. Because of the involvement of high contrast in edge detection, contrast changes result in different sets of edges. Thus, again, edge detectors cannot lead to a general shape comparison method.

2.2. The occlusion argument

In the preceding section, we have reduced an image, from the shape parsing viewpoint, to the set of all its level lines, otherwise called *topographic map*. The topographic map is in no way the ultimate description of shapes. When a shape A partially occludes a shape B , the level lines of the resulting image are a concatenation of the level lines belonging to A and to B . Thus, the Jordan curves of the topographic map are not simple shape elements and must be further decomposed: a segmentation of them into their parts belonging to different objects is requested. We can summarize the above discussion by a definition of shape elements.

Definition 1. We call *linelet* of an image u (or *shape element*) any piece of level line of any level line of the image.

Linelets are the most elementary shape elements we can rely on. We shall also give a tentative “blind” definition of shapes which relies on Gestalt theory. In the definition, we must definitely strengthen the invariance argument: planar objects are easily recognized under Euclidean, affine or projective transforms, or even more general deformations. In this paper, we shall only consider Euclidean or affine transforms. The notion of shape depends upon its invariance group.

Definition 2. We call *Euclidean* (resp. *affine*) *shape of an image* any *linelet*, or any *group of linelets*, which can be observed at least twice in the image, up to a *similarity* (resp. an *affine map*). We call *Euclidean* (resp. *affine*) *shape of a set of digital images* any *linelet* or *group of linelets* which can be retrieved in several images, up to *similarities* (resp. up to *affine maps*).

This second definition needs a distance measuring whether two shapes are the same or not (up to the allowed deformations and perturbations). We shall address this more technical question in detail in the next section. In the following, we denote by

$$d(X, Y) = \max \left(\sup_{x \in X} \inf_{y \in Y} \|x - y\|, \sup_{y \in Y} \inf_{x \in X} \|x - y\| \right)$$

the Hausdorff distance of two sets X and Y in Euclidean space.

Definition 3. We call *Euclidean* (resp. *affine*) *distance of two linelets* C and D ,

$$\delta(C, D) = \inf_{A \in G(\mathbb{R}^2)} d(AC, D),$$

where G is the group of similarities (or of affine maps) and D is usually normalized (diameter = 1 in the case of similarities, area = 1 in the case of the affine distance).

The scope of the normalization is of course to get a calibration of distances. Otherwise, small shapes would be closer to each other than large ones.

3. SHAPE SCALE SPACE

In the preceding section, we have given a definition of shape which leads us back to the question: how can we say that two shapes are equal, up to certain perturbations and a group of deformations? This question has actually been extensively treated, and in a very original way, in the computer vision literature. Authors have generally agreed to start with shapes defined as chained contours (Jordan curves), no matter how these curves have been extracted from the digital image. Because of the nature of digital (and biological) images, two identical contour curves undergo perturbations and deformations which make them look different enough in different images. This difference follows from different poses (inducing different deformations), different illuminations (inducing contrast changes), different resolution and perturbations due to the physical nature of the sensor, inducing blur, noise, quantization and sampling. The elimination of these small perturbations is a preliminary step to an accurate and robust comparison of the shapes. Now, the amount of perturbations can vary a lot from an image to another and the actual size of the objects matters too. A shape seen at long distance is of course more blurry and noisy than one in front (its signal-to-noise ratio is lower). The solution to this problem exists under the name of *scale space*. It consists in performing an *invariant smoothing* of the shape elements (understood as Jordan curves). By invariant we mean that the smoothing must commute with the main large perturbations: contrast changes and projective deformations. The first commutation is ensured by taking as shape contours all level lines of the digital image.

The main developments of Scale Space theory in the past ten years involve invariance arguments: indeed, a scale space will be useful to compute invariant information only if it is itself invariant. Let us summarize a

series of arguments given (*e.g.*) in [2]: a scale space computing contrast invariant information must in fact deal directly with the image level lines; in order to be local (not dependent upon occlusions), it must be in fact a partial differential equation. In order to be a smoothing, this PDE must be of a parabolic kind. Then, the further affine invariance requirement and the invariance with respect to reverse contrast (if we want a self-dual operator, in the morphomathematical terminology [34]) lead to a single PDE,

$$\frac{\partial u}{\partial t} = |Du| \text{curv}(u)^{\frac{1}{3}}, \tag{1}$$

where Du denotes the gradient of the image, $\text{curv}(u)$ the curvature of the level line, t denotes the scale parameter and the power $\frac{1}{3}$ is signed, *i.e.* $s^{\frac{1}{3}} = \text{sign}(s)|s|^{\frac{1}{3}}$. This equation is equivalent to the “affine curve shortening” [33] of all of the level lines of the image, given by the equation

$$\frac{\partial x}{\partial t} = |\text{Curv}(x)|^{\frac{1}{3}} \vec{n}, \tag{2}$$

where x denotes a point of a level line, $\text{Curv}(x)$ its curvature and \vec{n} the signed normal to the curve, always pointing towards the concavity.

This equation is the only possible smoothing under the invariance requirements mentioned above. This gives a helpless bottleneck to the local shape recognition problem, since it is easily checked [2] that no further invariance requirement is possible: the equation (1) is the only affine invariant local contrast invariant smoothing. In particular, despite some interesting attempts [13], there is no practical way to define a fully projective invariant local smoothing. Dibos [10], however, proposed a new invariance requirement, consistent with the displacement of a pinhole camera, the “registration group”. This group has six parameters like the planar affine group. Both are subgroups of the planar projective group with eight parameters.

The use of curvature-based smoothing for shape analysis is not new. See [4] for a founding paper, references [1, 26] and [12] for the definition of multiscale curvature which is similarity invariant, but not affine invariant.

There is another way to give an account of affine invariant smoothing, which is quite useful for the design of a fast algorithm. Using the mathematical morphology formalism [23, 34], one can define first an affine distance and then affine set erosions and dilations. Alternating small scale dilations and erosions yields a Chernouf formula for approximating equation (1). See Moisan [25].

Angenent *et al.* [3] proved that the affine shortening equation (2) indeed smoothes any Jordan curve and eventually shrinks it to an elliptic shape. Before this full reduction has been performed, the curve loses progressively its inflexion points and becomes simpler and simpler. The smoothing commutes with any affine transform with determinant 1. The “most invariant” shape smoothing being unique, we can use it as the standard tool to define affine invariant distances between shapes.

Consider two shapes C and D and their smoothed versions at scales s and t , $C(s)$ and $D(t)$. Under the assumption that AC is close to D for some affine map A , and according to the scale compatibility relation, we must have $AC((\det(A)^{\frac{2}{3}}t)$ close to $D(t)$ [14]. Indeed, $C(t)$ is solution of equation (2) if and only if $AC(\det(A)^{\frac{2}{3}}t)$ is. This leads us to a generalized definition of similarity between shapes.

Definition 4. We call scale-space distance of two shapes C and D the number

$$\Delta(C, D, t) = \inf_{A \in GL(2), t \in \mathbb{R}^+} d(AC((\det A)^{\frac{2}{3}}t), D(t)).$$

In this definition, d denotes the Hausdorff distance. The distance Δ has to be t -dependent: when $t \rightarrow +\infty$, all shapes tend to become equal. Thus, in order to decide about similarity of shapes, we must ask $\Delta(t)$ to be smaller and smaller when t increases.

4. EXPERIMENTS ON SHAPE

In the preceding sections, we have given a fairly general geometric definition of shape as any part of the topographic map which remains constant. We have justified this definition as the most general possible according to the local contrast invariance argument and the occlusion invariance argument. A comparison of all parts of the topographic map with all parts of another topographic map leads to combinatorial blow-up. In order to avoid it, the experiments will compare long enough level lines. The restriction that the shapes have long enough level lines is, however, only slightly restrictive and amounts to an average image quality requirement.

4.1. Summary of the algorithm computing the shapes in a set of images

We have defined a *shape element*, or *linelet* of an image as any piece of a level line. In order to compare this linelet to another, up to an Euclidean or affine transform, a normalization is performed. The normalized piece of linelet is discretized and the coordinates of the discretization points are used to form an array of points in a fixed reference system. This array of points is an Euclidean (resp. affine) unique representative of the shape, which can then be compared (for the sup norm) to other ones in the same or another image.

Given any two images, the registration algorithm, of which a detailed account can be found in [19] and [20], consists of four steps:

1. extraction of all the level lines for each digital image;
2. filtering of the extracted level lines at a fixed scale, large enough to remove digitization effects;
3. local encoding of all pieces of level lines or linelets after Euclidean (resp. affine normalization);
4. systematic comparison of codes of an image to the codes of another (or the same) image.

In order to perform a multi-scale strategy, these four steps should be repeated at several scales of filtering, and the final result should integrate the partial results obtained at each scale. In all the experiments we shall show here, however, the scale of affine smoothing is small, just enough to eliminate usual noise and sampling perturbations in a standard digital image. Also, the allowed zoom factor never exceeds 4 between two compared linelets.

Since our scope here is to stress the very definition of shape, we do not give here any implementation detail. The main novelty of the proposed method with respect to other existing shape comparison methods is the extraction of all the shape elements (completeness of the method) and the local encoding technique for the level lines. The extraction process generates a data structure (tree) of level lines similar to the one described in [28] and [21]. The simplification of this tree can be performed by using the method proposed in [9]. The filtering technique, applied directly to the extracted curves, is due to Moisan [25].

4.2. Experimental results: Shapes in a single image

The scope of the experiments simply is to display shapes, according to our definition. Surprisingly enough, most images of our natural environment contain several instances of objects: houses have windows and tiles, trees have leaves; people have two eyes... Characters, tissues, and most documents, posters, etc. present again and again variations on the same geometric figures, etc. Thus, most images are likely to define several shapes as repeated objects. We shall omit in the following to display the most repeated shape, namely the straight line. We shall only display the repeated linelets which have a (relatively) complex structure, with an angle total variation larger than π .

Our first experiment shows a poster of the Spanish Mathematical Association advertising mathematics which we have photographed with a standard low quality web camera. The lower central rectangle is an image with some eggs.

In the first row of the figure, one can see the photograph of the original poster (left) and a detail of the eggs image (right). On the second row, left, we have shown (in white) a level set (level 128) of the eggs image. The set of all level sets is a contrast invariant representation of shape information in an image proposed in Mathematical Morphology. In the central image of this row, we show a quantized topographic map of the eggs image, namely all level lines for levels which are multiple of 20. We finally show the blind shapes (right) of this

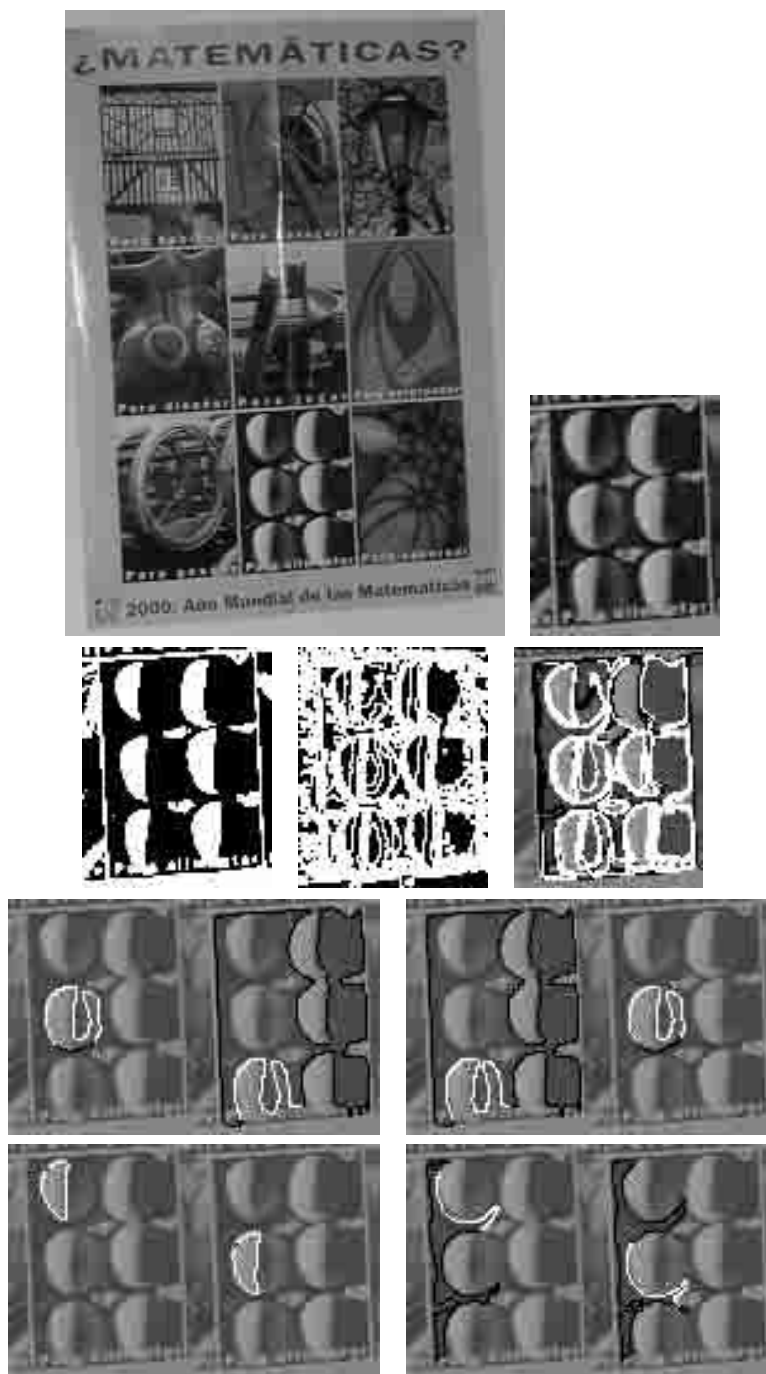


FIGURE 1. Experiment 1. First row, photograph of the original poster (left) and detail of the eggs image (right). Second row, level set (level 128) of the eggs image (left), its topographic map (quantization factor of the grey levels = 20) (center) and its blind shapes (right). Third and fourth rows, some instances of corresponding shapes. (White lines display the linelets, black lines show the remainder of the level line which contains the matching linelet).

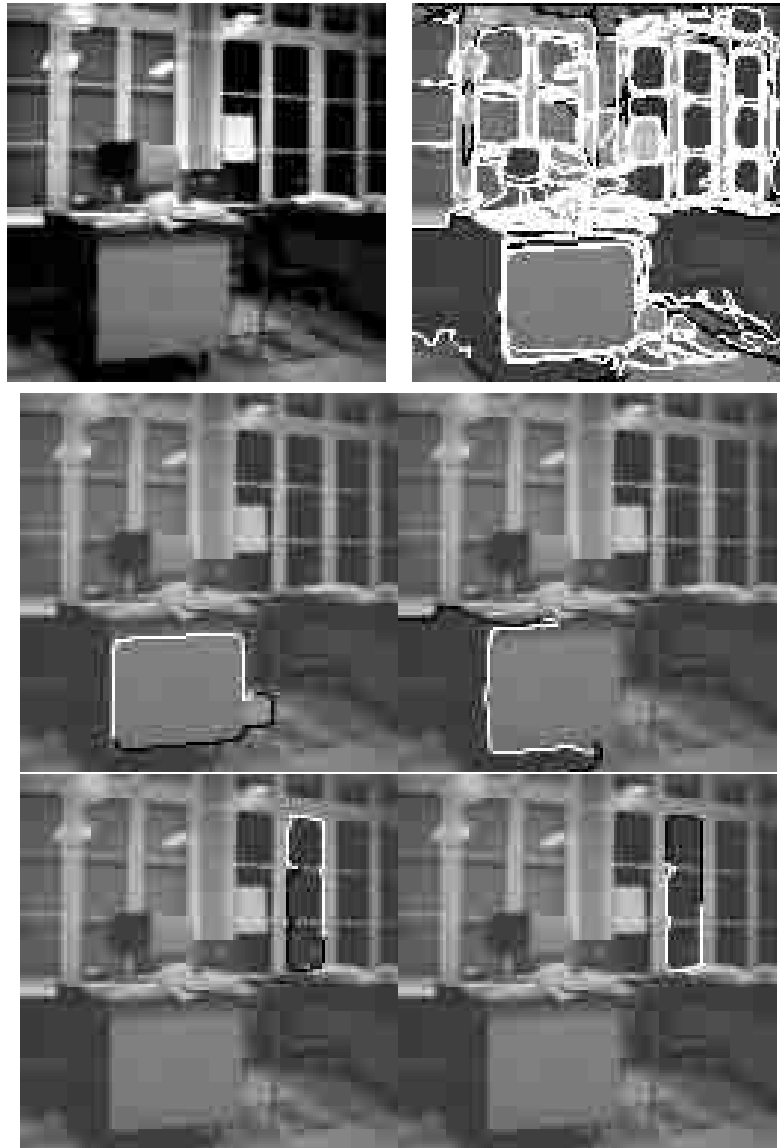


FIGURE 2. Experiment 2. First row, original image (left) and its blind shapes (right). Second and third rows, some instances of corresponding shapes. In the second row one sees that the affine self-similarity of a parallelogram is automatically detected by the algorithm.

image. The third and fourth rows are dedicated to some instances of pairs of corresponding shapes. (White lines display the linelets, black lines show the remainder of the level line which contains the matching linelet.) We see that, in an image made of repeated objects, the “blind shapes” according to our definition give a good account of the whole picture. In continuation, we display pairs of matching linelets in the eggs image. Those are the “blind shapes” according to our definition. They are typically the shapes which can be stored in a database and possibly be manually associated a name such as “eggs”.

Experiment 2 follows exactly the same format; this is a typical indoors scene, endowed with many rectangles seen projectively as rough parallelograms. Thus, they are similar for an affine invariant shape similarity measure



FIGURE 3. Experiment 3. Blind shape definition in a pair of images with a strong projective deformation. Both images have been taken by a camera with different illumination and pose. First row, original images. Second row, corresponding shapes between the images.

and also are self-similar, up to a circular permutation of the vertices. The second row shows that this self-similarity is detected.

In Experiment 3, we check which shapes can be recovered when we compare an image taken from very different poses and illuminations. For the rest, the format of the experiment is the same: in white, all “blind shapes”, *i.e.* linelets common to both images.

5. CONCLUSION

In this note, we have proposed an empirical definition of shape based only on two obvious principles: invariance to pose and illumination and repetition. We call this definition “blind shape” in that it does not correspond to any naming or learning process. We have shown that an algorithm exploring systematically all level lines can deliver automatically all blind shapes in an image or in several ones. This definition is merely geometric and appears as a necessary, automatic layer in any artificial intelligence or machine learning strategy.

The first author gratefully acknowledges partial support by CICYT project, reference TIC99-0266. The second author acknowledges partial support by CNES (Centre National d’Études Spatiales). The authors acknowledge partial support by US Navy, Office of Naval Research.

REFERENCES

- [1] S. Abbasi and F. Mokhtarian, Retrieval of similar shapes under affine transformation, in *Proc. International Conference on Visual Information Systems*. Amsterdam, The Netherlands (1999) 566-574.
- [2] L. Alvarez, F. Guichard, P.-L. Lions and J.M. Morel, Axioms and fundamental equations of image processing: Multiscale analysis and P.D.E. *Arch. Rational Mech. Anal.* **16** (1993) 200-257.
- [3] S. Angenent, G. Sapiro and A. Tannenbaum, On the affine heat flow for nonconvex curves. *J. Amer. Math. Soc.* (1998).
- [4] H. Asada and M. Brady, The curvature primal sketch. *PAMI* **8** (1986) 2-14.
- [5] L.G. Brown, A survey of image registration techniques. *ACM Comput. Surveys* **24** (1992) 325-376.

- [6] V. Caselles, B. Coll and J.M. Morel, Topographic maps and local contrast changes in natural images. *Int. J. Comput. Vision* **33** (1999) 5-27.
- [7] V. Caselles, B. Coll and J.M. Morel, Geometry and color in natural images. *J. Math. Imaging Vision* (2002).
- [8] T. Cohignac, C. Lopez and J.M. Morel, Integral and local affine invariant parameter and application to shape recognition, in *ICPR94* (1994) A164-A168.
- [9] A. Desolneux, L. Moisan and J.M. Morel, Edge detection by Helmholtz principle. *J. Math. Imaging Vision* (to appear).
- [10] F. Dibos, *From the projective group to the registration group: A new model*. Preprint (2000).
- [11] R.O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis*. Wiley (1973).
- [12] G. Dudek and J.K. Tsotsos, Shape representation and recognition from multiscale curvature. *CVIU* **2** (1997) 170-189.
- [13] O. Faugeras and R. Keriven, Some recent results on the projective evolution of 2d curves, in *Proc. IEEE International Conference on Image Processing*. Washington DC (1995) 13-16.
- [14] F. Guichard and J.M. Morel, *Image iterative smoothing and P.D.E.'s* (in preparation).
- [15] R.K. Hu, Visual pattern recognition by moments invariants. *IEEE Trans. Inform. Theor.* (1962) 179-187.
- [16] G. Kanizsa, Organization in vision: Essays on gestalt perception, in *Praeger* (1979).
- [17] A. Krzyzak, S.Y. Leung and C.Y. Suen, Reconstruction of two-dimensional patterns from Fourier descriptors. *MVA* **2** (1989) 123-140.
- [18] C.C. Lin and R. Chellappa, Classification of partial 2-d shapes using fourier descriptors, in *CVPR86* (1986) 344-350.
- [19] J.L. Lisani, *Comparison automatique d'images par leurs formes*, Ph.D. Dissertation. Université Paris-Dauphine (2001).
- [20] J.L. Lisani, L. Moisan, P. Monasse and J.M. Morel, Planar shapes in digital images. *MAMS* (submitted).
- [21] J.L. Lisani, P. Monasse and L. Rudin, Fast shape extraction and applications. *PAMI* (submitted).
- [22] D. Marr and E.C. Hildreth, Theory of edge detection. *Proc. Roy. Soc. London Ser. A* **207** (1980) 187-217.
- [23] G. Matheron, *Random Sets and Integral Geometry*. John Wiley, NY (1975).
- [24] W. Metzger, *Gesetze des Sehens*. Waldemar Kramer (1975).
- [25] L. Moisan, Affine plane curve evolution: A fully consistent scheme. *IEEE Trans. Image Process.* **7** (1998) 411-420.
- [26] F. Mokhtarian and A.K. Mackworth, A theory of multiscale, curvature-based shape representation for planar curves. *PAMI* **14** (1992) 789-805.
- [27] P. Monasse, Contrast invariant image registration, in *Proc. of International Conference on Acoustics, Speech and Signal Process.*, Vol. 6. Phoenix, Arizona (1999) 3221-3224.
- [28] P. Monasse and F. Guichard, Fast computation of a contrast-invariant image representation. *IEEE Trans. Image Processing* **9** (2000) 860-872.
- [29] M. Okutomi and T. Kanade, A locally adaptive window for signal matching. *Int. J. Computer Vision* **7** (1992) 143-162.
- [30] E. Persoon and K.S. Fu, Shape discrimination using fourier descriptors. *SMC* **7** (1977) 170-179.
- [31] T.H. Reiss, *Recognizing Planar Objects Using Invariant Image Features*. Springer Verlag, *Lecture Notes in Comput. Sci.* **676** (1993).
- [32] W.J. Rucklidge, Efficiently locating objects using the Hausdorff distance. *Int. J. Computer Vision* **24** (1997) 251-270.
- [33] G. Sapiro and A. Tannenbaum, Affine invariant scale-space. *Int. J. Computer Vision* **11** (1993) 25-44.
- [34] J. Serra, *Image Analysis and Mathematical Morphology*. Academic Press, New York (1982).
- [35] C.H. Teh and Chin R, On image analysis by the method of moments. *IEEE Trans. Pattern Anal. Machine Intelligence* **10** (1998).