In *Visual Form 2001*, C. Arcelli, L.P. Cordella, G. Sanniti di Baja eds., Lecture Notes in Computer Science, Springer Verlag, pp. 285-294, 2001.

Global topological properties of images derived from local curvature features

Erhardt Barth, Mario Ferraro[&], and Christoph Zetzsche[#]

Institute for Signal Processing
University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany
[&] Dipartimento di Fisica Sperimentale and INFM
via Giuria 1, Torino, Italy
[#]Institut für Medizinische Psychologie, Goethestr. 31, 80336 München, Germany
barth@isip.mu-luebeck.de, ferraro@to.infn.it,
chris@imp.med.uni-muenchen.de
http://www.isip.mu-luebeck.de

Abstract. In this paper we show that all images are topologically equivalent. Nevertheless, one can define useful pseudo-topological properties that are related to what is usually referred to as topological perception. The computation of such properties involves low-level structures, which correspond to end-stopped and dot-responsive visual neurons. Our results contradict the common belief that the ability for perceiving topological properties must involve higher-order, cognitive processes.

Keywords: Topology, Euler number, closure, curvature, visual perception

1 Introduction

A fundamental issue of image analysis and understanding is the mathematical description, or representation, of the input data. The basic representation of an image and indeed the closest to the physical process of vision is that of a graph of a function $I : (x, y) \rightarrow I(x, y)$ from a subset U of R^2 to R; here U is the retinal plane and $I(x, y)$ is the light intensity at the point $(x, y) \in U$. From a geometrical point of view, the graph of I is a Monge patch, that is a surface L of the form $L = \{x, y, I(x, y)\}$. Alternatively, L can be considered as the image of the mapping $\phi : (x, y) \rightarrow (x, y, I(x, y))$. Then, the process of image formation can be seen as a mapping from the surface M of a physical object to the Monge patch L representing the image. Let S be the surface of an object in R^3. Consider a coordinate system $\{x, y, z\}$, whose origin coincides with the position of the viewer and let $\{x, y\}$ be the image plane. The visible part M of S can be given a Monge-patch representation $M = (x, y, f(x, y))$, where $z = f(x, y)$ is the distance of the point $r = (x, y)$ on the image plane of the observer to $s = (x, y, z)$ in M. Note that, if orthographic projection is assumed, f and I share the same domain U and there is a one-to-one correspondence between $p = (x_0, y_0, f(x_0, y_0)) \in M$ and $q = (x_0, y_0, I(x_0, y_0)) \in L$ [8]. The

extension to the case of perspective projection is immediate. If a surface M is transformed by a continuous transformation into a surface M', a new image L' will be generated by M'; the relation between the transformation of the image and the transformation of the underlying object has been made precise in [15].

The representation of images as surfaces has two advantages: first it is close to the original data structure and does not require any high level process to be generated, second it allows one to use the technical machinery of geometry to investigate it. In particular the condition of differentiability allows to make use of a very powerful theorem of differential geometry, the Gauss-Bonnet theorem, which provides a link between global and local properties of surfaces. Surfaces can be given a global classification based on the notion of topological invariants. A property of a surface is called a topological invariant if it is invariant under homeomorphism, that is under an one-to-one continuous transformation that has a continuous inverse. Two surfaces are said to be topologically equivalent if they can be transformed one in the other by a homeomorphism. For instance, the number of holes in a surface is a topological invariant. In particular here we want to investigate, how the topological properties of object surfaces are reflected in the images. It will be shown in the next section that all images are topologically equivalent, i.e. any two images can be transformed one into the other by means of an homeomorphism. From this result it follows that the topological properties of an object's surface are not intrinsic properties of its image.

We shall show how topological properties of the object underlying the images can be found by using $2D$-operators, i.e. operators whose output is different from zero only in case of $2D$-features such as corners, line-ends, curved lines and edges. These operators can be associated to the activity of nonlinear end-stopped neurons. Mechanisms possibly underlying the activity of such neurons have been investigated by a few authors, e.g. [7, 9, 14, 16]. A general theory for end-stopping and $2D$-operators, however, is still missing, but attempts to identify the basic ingredients for such a theory have been made [16, 17].

Finally, we will show how $2D$-operators can be used to provide an alternative explanation for some experimental findings [4, 5], which have suggested that the human visual system might be quite sensitive to global, "topological" characteristics of images.

2 The Gauss-Bonnet theorem

The Gauss-Bonnet theorem is one of the most important theorems in differential geometry, in that it provides a remarkable relation between the topology of a surface and the integral of its Gaussian curvature.

Let R be a compact region (e.g. a Monge patch) whose boundary ∂R is the finite union of simple, closed, piecewise regular curves C_i. Consider a polygonal decomposition of R, that is a collection of polygonal patches that cover R in a way such that if any two overlap, they do so in either a single common vertex or a single common edge [12].

Thus a polygonal decomposition $\mathcal{D}$ carries with it not only the polygonal patches, called faces, but also the vertices and edges of these patches. Suppose $\mathcal{D}$ has f faces, e edges and v vertices, then $\chi = f - e + v$ is the Euler-Poincare' characteristic of R and it is the same for all polygonal decompositions of R [12]. The Euler-Poincare characteristic can be extended, in a natural way, to regular, compact surfaces. In Fig. 1 some examples of closed surfaces with different values of χ are shown. Two compact surfaces are topologically equivalent if and

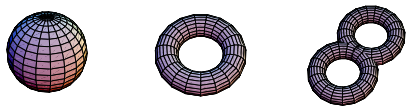

Fig. 1. Surfaces with different Euler-Poincare' characteristics ($\chi = 2, 0, -2$ from left to right).

only if they have the same Euler-Poincare' characteristic χ [12, 6]. Here, we are interested in Monge patches, that represent both the visible part of surfaces and their images. If the Monge patch M is a simple region, that is is homeomorphic to an hemisphere, $\chi = 1$ [6] , if it has an hole $\chi = 0$, and in general

$$\chi = (1 - n_{holes}). \tag{1}$$

Note that if, instead of regions, we consider regular surfaces, equation 1 becomes $\chi = 2(1 - n_{holes})$. The definition of χ holds for connected surfaces; we shall now extend the definition of χ to the case of not connected surfaces. Suppose there are n object in the scene, the global Euler-Poincare' characteristic χ_T is then simply $\chi_T = \sum_j^n \chi_j$.

Consider a region R. The Euler-Poincare' characteristic is related to the curvature of R by the celebrated Gauss-Bonnet formula

$$\iint_R K dA + \sum_i \int_{C_i} k_g ds + \sum_i \theta_i = 2\pi\chi, \tag{2}$$

here dA is the infinitesimal area element, C_i are regular curves forming the boundary ∂R of R, k_g is the geodetic curvature computed along the curve C_i and θ_i are the external angles at the vertices of ∂R.

If S is a compact orientable surface then $\int\int_S K dA = 2\pi\chi(S)$ [12, 6].

This is a striking result: how is it possible that, when integrating a local property of a surface, we obtain global topological invariants of that surface? Let us consider the surface of a sphere S^2 in R^3. For a given radius r, the Gaussian curvature is $K = 1/r^2$ and $\int\int_{S^2} K dA = 4\pi$. Note that this result does not depend on r, since as r increases K decreases but the area of the surface increases. More importantly, the result is the same for all surfaces topologically

equivalent to a sphere. Suppose S^2 is transformed into a surface S by deforming the sphere with a dent. In this case the area increases, and the elliptic part of the dent gives additional positive values. At the same time, however, new negative values of K are produced at the hyperbolic regions of the dent, and, as a consequence, positive and negative curvatures cancel out and the total curvature remains constant.

3 All images are topologically equivalent

In this section we give a formal proof that all images are topologically equivalent. Topological equivalence of two images implies that, given any pair of images with representation $L = (x, y, I(x, y))$ and $N = (x, y, J(x, y))$, there exists an homeomorphism taking one into the other. First we prove the following

Lemma 1 *Let I be a function from U to R and let $\phi : (x, y) \to (x, y, I(x, y))$ from U to R^3. If I is continuous then ϕ is an homeomorphism.*

Proof. The map ϕ is one-to-one and can be written as $\phi = i \times I$ where $\times$ is the Cartesian product and i is the identity map of U onto itself, which is obviously continuous. Since I is continuous by hypothesis, then ϕ is continuous [6]. The inverse of ϕ, is nothing else than the orthographic projection from M to U, which is continuous [6] and the assertion follows.

From the lemma it follows that

Proposition 1 *Let L and N be images defined as the graph of continuous functions I and J respectively; then there exists an homeomorphism $h : L \to N$.*

Proof. Consider the maps $\phi : (x, y) \to (x, y, I(x, y))$ from U to L and $\psi : (x, y) \to (x, y, J(x, y))$ from U to N and define $h = \psi \circ \phi^{-1}$, which is a map from L to N. The map h is an homeomorphism being the composition of two homeomorphisms ϕ and ψ.

If f is supposed to be smooth, that is to have continuous partial derivatives of any order in a open set U in R^2, then the surface is regular [6], and then it is easy to prove that all surfaces, which are the graph of differentiable functions, are diffeomorphically equivalent; that is to say that for any pair of surfaces there exists a bijective map, smooth with its inverse, taking one surface onto the other.

From proposition 1 it follows that all images must have the same Euler-Poincare' characteristic χ, that can now be computed by making use of the Gauss-Bonnet theorem.

The Gaussian curvature of a Monge patch $L = \{x, y, I(x, y)\}$ is given by [6]

$K = \mid H_I \mid \left(1 + (\partial I/\partial x)^2 + (\partial I/\partial y)^2\right)^{-2}$ where $\mid H_I \mid$ is the determinant of the Hessian matrix H_I of the function I.

Let now $V \subset R^2$ be a disk of radius r with boundary C; V is a flat surface, hence $\int\int_V K dA = 0$ and $\sum_i \theta_i = 0$ because there are no vertices. In this case the geodetic curvature along C is equal to the curvature $k = 1/r$ of C and it follows

that $\int_C k_g ds = 2\pi$. Therefore $\chi = 1$, and, since χ is a topological invariant, it must be the same for all images.

It must be pointed out that the result applies only to Monge patches, and hence not to any regular surface in R^3; however we are interested in images, which indeed are graphs of functions, and hence Monge patches.

In our proof we assumed image intensity I to be a continuous function. This is a common assumption justified by the observation that most images are band limited due to the imaging system. In human vision there is a good match between the band limitation and the sampling density.

We have seen here that all images are topologically equivalent. Of course this is is not true for the objects that generate different images. The surfaces of these objects may have different topological properties, e.g. for a sphere and a torus $\chi = 2$ and $\chi = 1$ respectively (see Fig. 1), and their visible parts have $\chi = 1$ and $\chi = 0$ respectively (see Eq. 1); however, their images have $\chi = 1$. Thus the topological properties of an object's surface cannot be determined as topological properties of the corresponding image surfaces. In other words, characteristics such as holes or discontinuities do not exist in the images *per se*, indeed there is no a priori reason to interpret dark or light blobs as holes. Nevertheless, one can do so and successful methods for estimating the Euler characteristic of binary images have been presented [13, 10].

4 Pseudo-topological properties of images

It has been mentioned before that there is some experimental evidence suggesting that the human visual system can discriminate on the basis of what appear to be different topological properties of the objects underlying the image. Then it is of interest to search for image properties, which reflect topological properties of the associated objects. We call such properties pseudo-topological properties of images, and we shall investigate, which kind of operators are appropriate to detect pseudo-topological image properties. To compute pseudo-topological properties of images by integrating local features as the outputs of some operator, then only 2D-operators, that capture the local curvature of the image, seem appropriate, even though not all $2D$ global operators will work in detecting pseudo-topological properties of images. Indeed let L be the geometrical representation of an image and suppose that its boundary C is a regular closed curve contained in a planar region such that, in C, $|\nabla I| = 0$. Then we have, as seen before,

$$\sum_i \theta_i + \sum_i \oint_{C_i} k_g ds = 2\pi. \tag{3}$$

From the Gauss-Bonnet formula $\int\int_V K dA + 2\pi = 2\pi\chi$, and, since $\chi = 1$, $\int\int_V K dA = 0$. If we now extend an image by a planar frame we can find a curve C such that Eq. (3) holds. Then, for all "framed" images the total curvature is equal to zero. For any practical purpose this also implies that any

deviation from zero will be a boundary effect. Therefore, a straightforward application of the Gauss-Bonnet theorem to image analysis cannot lead to useful image properties. We will make no attempt to develop a general theory about the invariance properties of integral 2D-operators. Instead, we will show how a specific 2D-operator, namely the clipped-eigenvalues (CEV) operator, can be used to compute pseudo-topological properties of images. This operator has been introduced as a model for dot-responsive cells in [16], and described in more detail in [1]. Nevertheless, it seems useful to understand in the present context how the operator can be derived from the expression for the Gaussian curvature K. The determinant of the Hessian of I can be written as

$$|H_I| = \frac{1}{4}\left(\frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}\right)^2 - \frac{1}{4}\left(\frac{\partial^2 I}{\partial x^2} - \frac{\partial^2 I}{\partial y^2}\right)^2 - \left(\frac{\partial^2 I}{\partial x \partial y}\right)^2, \tag{4}$$

that is,

$$|H_I| = \left(\nabla^2 I\right)^2 - \epsilon^2, \tag{5}$$

where $\nabla^2 I$ is the Laplacian on the intensity I and ϵ is the eccentricity; they determine the eigenvalues of H_I through the formula $\lambda_{1,2} = \nabla^2 I \pm \epsilon^2$. The operation of clipping is defined as $\lambda^+ = Max(0, \lambda)$ and $\lambda^- = Min(0, \lambda)$. The CEV operator is then $CEV(I) = \lambda_2^+(I) - \lambda_1^-(I)$. Note that in case of isotropic patches, where $\epsilon = 0$, $CEV(I) = \nabla^2 I$. But when a pattern becomes elongated, CEV will be less than $\nabla^2 I$ and will become zero for straight patterns. The local operator CEV yields a global measure $< CEV >$ defined as the average of CEV on the whole image.

The main difference between the Gaussian curvature and the operator CEV is that the latter is zero for hyperbolic surface patches. Using the clipping operation one obtains different signs for positive- and negative-elliptic patches, and only patches with absolute CEV values above a small threshold contribute to $< CEV >$. Fig. 2 presents some examples of the action of CEV on different images and the corresponding values of $< CEV >$ are also shown. Note that $| < CEV > |$ is, within a very small error due to the numerical approximation, equal to $|\chi_T|$, the total Euler-Poincare' characteristic of the visible parts of surfaces in the scene.

The $< CEV >$ measure exhibits the pseudo-topological invariance illustrated above as long as the patterns have the same contrast. A contrast independent version of the CEV operator can be defined as $CEV_N = \frac{\left(\lambda_2^+(I) - \lambda_1^-(I)\right)}{(c + (\partial I/\partial x)^2 + (\partial I/\partial y)^2)^{1/2}}$, where c is a small constant. With this measure we have obtained results where $< CEV_N >$ varied with less than 1% for patterns with different contrasts.

Pseudo-topological invariance is also limited by the scale of the operators and the size and shape of the patterns that are involved. We should mention here, that before computing the partial derivatives, the images are low-pass filtered with a Gaussian filter, which defines the scale of the CEV operator (that can be evaluated on multiple scales). However, any image can be zoomed (by nearest-neighbor interpolation) such that patterns are approximated by saw-tooth contours (like the two tilted polygons in the third row of Fig. 2). In this

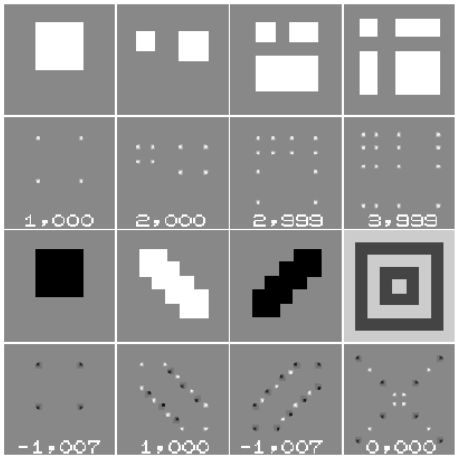

Fig. 2. Responses of the CEV operator (second and fourth row) to 8 different patterns (first and third row). In addition, the mean values of $< CEV >$ (normalized to the first image) are given.

case the scale on which the CEV operator is computed can be finer than the highest frequencies of the original patterns. On such zoomed images, we can obtain the above pseudo-topological invariance independent of the shape of the patterns.

As concerns simulations of visual functions, however, the issue of strong invariance is less relevant. More important is whether we can predict the experimental results with reasonable model assumptions.

5 Simulations of experimental data

Due to Minsky and Papert [11] it is a widespread belief that topological properties of images can neither be easily computed by simple neural networks, nor easily perceived by human observers. In a series of experiments, Chen has shown that subjects are sensitive to "topological" properties of images, and he has interpreted his results as being a challenge to computational approaches to vision. For example in [4] he has shown that if two images are presented briefly (5 ms), subjects discriminate better between images of a disk and a torus, than they do in case of images of a disk and a square or a disk and a triangle, respectively. Note that the objects disk, square, and triangle are topologically equivalent, $\chi = 2$, whereas is case of a torus $\chi = 0$, compare Eq. (1) and Fig. 1.

Further indications of this kind of performance, that Chen attributed to some topological process, have been found in reaction-time experiments [5] where subjects were as fast (750 ms on average) in finding the quadrant containing a "closed" pattern (a triangle), with the 3 other quadrants containing an "open"

pattern (an arrow), as they were in finding an empty quadrant (with the other 3 quadrants containing a square) - see Fig. 4.

Proposition 1, on the other hand, demonstrates that the image does not directly exhibit the topological properties of the underlying surfaces and that some type of further processing is needed, which we have attributed to the action of the CEV operator. To simulate the results obtained by Chen, the image shown at the left of Fig. 3 was used as an input, to which the CEV operator was applied with a result shown in the middle of Fig. 3. The final results are displayed on the right of Fig. 3. Here we have not computed the global mean values $< CEV >$ but have integrated the local CEV values by low-pass filtering.

Therefore the intensity map shown on the right of Fig. 3 can be interpreted as a local estimate of $< CEV >$, denoted by CEV_{LP}, that varies with (x, y) and depends on the chosen scale. Obviously, if CEV_{LP} were the representation, which subjects use for discrimination, the difference between the ring and the disc would be larger than the differences between the disc and the rectangle and triangle (as found in the experiment). The existence of an end-stopped CEV-like representation is well motivated by neurophysiological and some psychophysical results [17]. The spatial filtering is common in many other models, e.g, of texture perception [3]. What it assumes is that the similarity metric involves some spatial integration. Of course, the CEV_{LP} representation depends on a few parameters, mainly the scale of CEV itself and of the low-pass filter. However, the point here is that it is easy to predict the experimental results with reasonable values of the parameters, and we found the predictions to be stable with respect to variations of the parameters. A more comprehensive analysis of how the pseudo-topological properties depend on the spatial scales of the underlying operations is beyond the scope of this paper. The results shown in Fig. 4 have been obtained

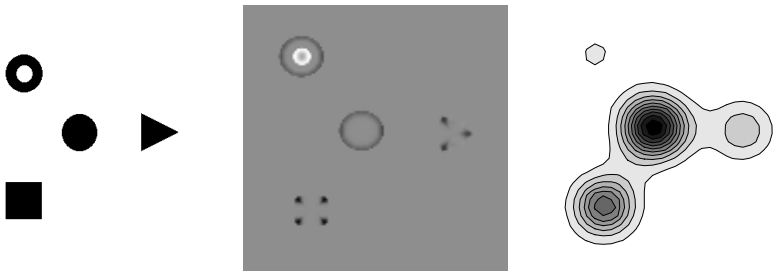

Fig. 3. Input image (left), output of the CEV operator (middle) and, on the right, low-pass filtered CEV operator (CEV_{LP}). The result on the right predicts the experimental findings that humans are more sensitive to the difference between a circle and a ring, than between a circle and a square or a triangle.

in a similar way. Here the CEV_{LP} result shown on the right illustrates the large difference between the triangle and the arrow in this representation.

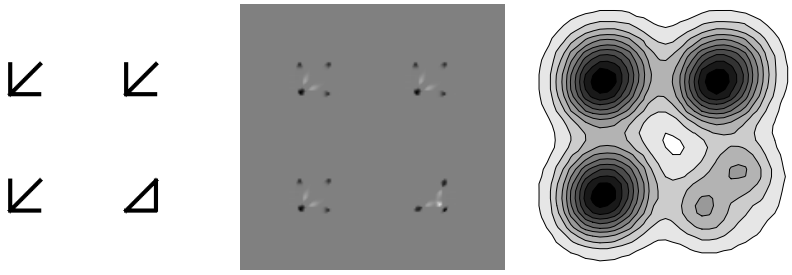

Fig. 4. Simulations as in Fig. 3 but for a different input. In this example results predict the large perceptual difference between the open and the closed shapes arrow and triangle.

6 Conclusions

We have shown here that all images are topologically equivalent. From this we can conclude that any "topological" properties of images depend on, and are restricted to, some additional abstractions or computational rules.

Chen's experiments reveal that the human visual system is sensitive to "topological" properties of the input patterns. Our simulations show that the results can be explained by assuming that the visual system evaluates integral values of specific curvature measures. These integral values can be seen as corresponding to activities of end-stopped, dot-responsive cells that are averaged over space. A possible interpretation is that in certain cases, e.g. at short-time presentations, the system evaluates integral values of an underlying, end-stopped representation. End-stopped neurons in cortical areas V1 and V2 of monkeys are oriented and more complex than dot-responsive cells. For simplicity, we have restricted our simulations to the CEV operator and have argued that the basic requirement for pseudo-topological sensitivity is that straight features are not represented. However, we have shown before that even the retinal output could be endstopped, depending on the dynamics of the input, and that the quasi-topological sensitivity is not limited to the use of the CEV operator [2].

The evaluation of integral values is assumed to be relevant to texture perception also. Indeed, we have been able to show that certain human performances in texture segmentation can be predicted by integrating the output of $2D$-operators in general and of the CEV operator in particular [3]. Thus, we are confident to be dealing with a rather general principle of visual processing.

7 Acknowledgment

This work is based on an earlier manuscript, which had been supported by a grant from the Deutsche Forschungsgemeinschaft (DFG-Re 337/7) to I. Rentschler and C. Z. We thank C. Mota and the reviewers for valuable comments.

References

1. E. Barth, T. Caelli, and C. Zetzsche. Image encoding, labelling and reconstruction from differential geometry. *CVGIP:GRAPHICAL MODELS AND IMAGE PROCESSING*, 55(6):428–446, 1993.
2. E. Barth and C. Zetzsche. Endstopped operators based on iterated nonlinear center-surround inhibition. In B. Rogowitz and T. Papathomas, editors, *Human Vision and Electronic Image Processing*, volume 3299 of *Proc. SPIE*, pages 67–78, Bellingham, WA, 1998.
3. E. Barth, C. Zetzsche, and I. Rentschler. Intrinsic two-dimensional features as textons. *J. Opt. Soc. Am. A*, 15(7):1723–1732, July 1998.
4. L. Chen. Topological structure in visual perception. *Science*, 218(12):699–700, 1982.
5. L. Chen. Topological perception: a challenge to computational approaches to vision. In e. a. P Pfeifer, editor, *Connectionism in perspective*, pages 317–329. Elsevier Science Publishers, North-Holland, 1989.
6. M. P. Do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1976.
7. A. Dobbins, S. W. Zucker, and M. S. Cynader. Endstopped neurons in the visual cortex as a substrate for calculating curvature. *Nature*, 329:438–41, 1987.
8. M. Ferraro. Local geometry of surfaces from shading analysis. *Journal Optical Society of America*, A 11:1575–1579, 1994.
9. J. J. Koenderink and W. Richards. Two-dimensional curvature operators. *J. Opt. Soc. Am. A*, 5(7):1136–1141, 1988.
10. C.-N. Lee and A. Rosenfeld. Computing the Euler number of a 3D image. Technical Report CAR-TR-205, CS-TR-1667, AFOSR-86-0092, Center for Automation Research, University of Maryland, 1986.
11. M. Minsky and S. Papert. *Perceptrons*. MIT Press, Cambridge MA, 1969.
12. B. O'Neill. *Elementary Differential Geometry*. Academic Press, San Diego CA, 1966.
13. A. Rosenfeld and A. C. Kak. *Digital Picture processing*. Academic Press, Orlando, FL, 1982.
14. H. R. Wilson and W. A. Richards. Mechanisms of contour curvature discrimination. *J Opt Soc Am A*, 6(1):106–15, 1989.
15. A. Yuille, M. Ferraro, and T. Zhang. Image warping for shape recovery and recognition. *Computer Vision and Image Understanding*, 72:351–359, 1998.
16. C. Zetzsche and E. Barth. Fundamental limits of linear filters in the visual processing of two-dimensional signals. *Vision Research*, 30:1111–1117, 1990.
17. C. Zetzsche, E. Barth, and B. Wegmann. The importance of intrinsically two-dimensional image features in biological vision and picture coding. In A. Watson, editor, *Digital images and human vision*, pages 109–138. MIT Press, Cambridge, MA, 1993.