Image Representation Using Affine Covariant Coordinates *

Jun Zhang

Department of Psychology, Cognition and Perception Programs,
University of Michigan, Ann Arbor, Michigan 48104-2994, USA

Abstract. To achieve affine-invariant image representation and shape recognition, one must rely on a set of affine *co*variant image coordinates. It is proposed that such coordinates be derived from the second derivatives (Hessian) of a grey-level image. The two eigen-directions of the image Hessian are everywhere orthogonal. Connecting corresponding direction at neighbouring points results in a smooth flow field. Proper parameterization by Lie bracket operation gives rise to two orthogonal flow fields that may serve as the coordinate bases for an arbitrary image. From its construction, this coordinating "net" covaries with affine transforms of the visual manifold. Topological deformation of the image shape can be concisely described as Lie group actions on these curvilinear coordinates.

Keywords: shape description, Lie transformation group, image Hessian, Lie bracket, image coordinates, gauge transformation, Gestalt image.

1 Introduction

A central issue in shape representation is that of its invariance. It is common sense that human object recognition is invariant under linear transformations of visual space that may involve translation, scaling, and, to a certain extent, rotation ("affine" transforms technically). In order to derive affine-invariant descriptors of image shape, one must look for a set of affine *co*variant coordinates of the visual manifold (i.e. coordinates that covary with affine space transforms) to counteract the consequence of an affine transform on shape descriptors. In previous studies [3, 1, 4, 6] shape invariance under a global and uniform translation, rotation, or scaling was achieved through a mapping of the visual space onto itself via the action of a corresponding Lie group. The generators of those Lie transformation groups are image-independent, that is, they do not involve the specific image under analysis.

There is yet another kind of shape invariance, that is, under moderate but arbitrary deformation of the visual space where shapes are defined. Similarity

* The author thanks the McDonnell-Pew Center for Cognitive Neuroscience at San Diego and Dr. Terrence J. Sejnowski for support.

judgements can be easily made on an object undergoing various degrees and manners of distortion. However, different image shapes may tolerate different patterns of distortion before visual recognition finally breaks down. In other words, shape invariance under this local and point-wise distortion is image-specific. To characterize such distortion (and therefore achieve invariance), the set of desired coordinates, which serve to re-partition the visual space, ought to be derived from a grey-level image. The image-dependent coordination of the space will not only enable invariant shape description under *global* affine transforms, but also allow convenient expressions for *local* symmetry at individual locations associated with relatively "harmless" deformation.

The intention here is to derive such coordinates for an arbitrary grey-level image. Noting that the two-dimensional visual manifold is the natural support of shape perception (and visual perception in general), it is important to understand first why the said manifold should be "coordinated" in one way as opposed to another. One immediate reason is that a good, image-driven coordination of visual manifold facilitates (or even enables) perceptual processing. It has previously been suggested [8] that visual perception has the mathematical structure of fibre bundles. The sensory representation of image attributes is described as constructing vector fields defined on the visual space (base manifold). Image segmentation is achieved by identifying intrinsically constant portions of the sensory vector fields or cross-section of the fibre bundle under a given connection on the base manifold. The connection for the tangent bundle of the visual manifold (i.e. the space under which image motion is processed) has been derived. The "good" coordinates or geodesics are given by the first-order image gradients. The metric tensor consistent with this interpretation was shown to equal the square of the second-order gradients (the Hessian) of a grey-level image:

$$g_{\mu\nu} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} , \tag{1}$$

where f_{xx}, f_{yy}, and f_{xy} are second-order spatial derivatives of an image function or grey-level intensity distribution $f(x, y)$. The metric tensor is, quite naturally, symmetric (with respect to its lower indices) and semi positive-definite (non-negative). The associated Riemann curvature tensor is identically zero, indicating that it is indeed possible to globally define image coordinates on the two-dimensional base manifold where image shape is to be defined. The task then is to find these image coordinates based on the above-mentioned geometrical framework of visual perception.

2 Establishing Image-based Coordinates

Let us start by diagonalizing the image Hessian and, according to (1), the metric tensor of the visual manifold (the latter is called "perceptual" metric to distinguish it from the trivial, Euclidean metric). The characteristic directions or eigenvectors are (with corresponding eigenvalues λ_1, λ_2)

$$\begin{cases} \mathbf{n}_1 = \cos\phi\, \mathbf{e}_1 + \sin\phi\, \mathbf{e}_2 \\ \mathbf{n}_2 = -\sin\phi\, \mathbf{e}_1 + \cos\phi\, \mathbf{e}_2 \end{cases} \tag{2}$$

where ϕ is the angle to be rotated with respect to a pre-chosen Cartesian coordinate basis $\mathbf{e}_1$, $\mathbf{e}_2$ such that

$$\begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} . \tag{3}$$

The second derivatives (Hessian) of an image function $f(x,y)$ can be explicitly related to λ_1, λ_2, and ϕ, after some rearranging:

$$\begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} = \begin{pmatrix} \lambda_1 \cos^2\phi + \lambda_2 \sin^2\phi & (\lambda_1 - \lambda_2)\sin\phi\cos\phi \\ (\lambda_1 - \lambda_2)\sin\phi\cos\phi & \lambda_1 \sin^2\phi + \lambda_2 \cos^2\phi \end{pmatrix} . \tag{4}$$

The directional derivatives (Lie derivatives) along $\mathbf{n}_1$, $\mathbf{n}_2$ are, respectively,

$$\frac{d}{dl_1} = \cos\phi \frac{\partial}{\partial x} + \sin\phi \frac{\partial}{\partial y} , \qquad \frac{d}{dl_2} = -\sin\phi \frac{\partial}{\partial x} + \cos\phi \frac{\partial}{\partial y} . \tag{5}$$

Since

$$\frac{\partial f_{xx}}{\partial y} = \frac{\partial f_{xy}}{\partial x} , \tag{6}$$

it follows that

$$\begin{aligned} &\left(\sin\phi \frac{d}{dl_1} + \cos\phi \frac{d}{dl_2}\right)(\lambda_1 \cos^2\phi + \lambda_2 \sin^2\phi) \\ &= \left(\cos\phi \frac{d}{dl_1} - \sin\phi \frac{d}{dl_2}\right)((\lambda_1 - \lambda_2)\sin\phi\cos\phi) \ ; \end{aligned} \tag{7}$$

Since

$$\frac{\partial f_{xy}}{\partial y} = \frac{\partial f_{yy}}{\partial x} , \tag{8}$$

it follows that

$$\begin{aligned} &\left(\sin\phi \frac{d}{dl_1} + \cos\phi \frac{d}{dl_2}\right)((\lambda_1 - \lambda_2)\sin\phi\cos\phi) \\ &= \left(\cos\phi \frac{d}{dl_1} - \sin\phi \frac{d}{dl_2}\right)(\lambda_1 \sin^2\phi + \lambda_2 \cos^2\phi) \ . \end{aligned} \tag{9}$$

Simplifying (7) and (9) yields

$$\frac{d\lambda_2}{dl_1}\sin\phi + \frac{d\lambda_1}{dl_2}\cos\phi + \frac{d\phi}{dl_1}(\lambda_2 - \lambda_1)\cos\phi + \frac{d\phi}{dl_2}(\lambda_2 - \lambda_1)\sin\phi = 0 \ , \tag{10}$$

$$\frac{d\lambda_2}{dl_1}\cos\phi - \frac{d\lambda_1}{dl_2}\sin\phi - \frac{d\phi}{dl_1}(\lambda_2 - \lambda_1)\sin\phi + \frac{d\phi}{dl_2}(\lambda_2 - \lambda_1)\cos\phi = 0 \ , \tag{11}$$

or, written compactly ($(\cdot,\cdot)^T$ denotes matrix transposition),

$$(\lambda_2 - \lambda_1)\begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}\left(\frac{d\phi}{dl_1}, \frac{d\phi}{dl_2}\right)^T + \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}\left(\frac{d\lambda_1}{dl_2}, \frac{d\lambda_2}{dl_1}\right)^T = \begin{pmatrix} 0 \\ 0 \end{pmatrix} . \tag{12}$$

Therefore

$$\frac{d\phi}{dl_1} = \frac{1}{\lambda_1 - \lambda_2}\frac{d\lambda_1}{dl_2}\,, \qquad \frac{d\phi}{dl_2} = \frac{1}{\lambda_1 - \lambda_2}\frac{d\lambda_2}{dl_1}\,. \tag{13}$$

Note that, although the three functions λ_1, λ_2, and ϕ are algebraically independent at any image point, they are analytically related to each other at a common neighbourhood, as indicated by (13). This is because the Hessian of any image function satisfies (6) and (8).

Now, at each image location, there is a pair of orthogonal unit vectors (tiny "needles") which represent eigen-directions of the image Hessian. Assuming they are continuous vector fields (i.e., for sufficiently smooth images, presumably after filtering), the corresponding eigenvectors at neighbouring locations may be connected to form two orthogonal flow fields, which can in turn act on the two-dimensional visual manifold. Either flow field will "fill up" a surface patch; together, they mesh into a new coordinating "net" for the underlying manifold. However, the eigenvector fields (5) do not themselves form coordinate bases. To see this, calculate their Lie bracket which expresses the commutativity of the two flow field actions (c.f. [5, pp. 42-49]):

$$\begin{aligned}\left[\frac{d}{dl_1},\frac{d}{dl_2}\right] &= \frac{d}{dl_1}\left(\frac{d}{dl_2}\right) - \frac{d}{dl_2}\left(\frac{d}{dl_1}\right)\\ &= \left(\cos\phi\frac{\partial}{\partial x} + \sin\phi\frac{\partial}{\partial y}\right)\left(-\sin\phi\frac{\partial}{\partial x} + \cos\phi\frac{\partial}{\partial y}\right)\\ &\quad - \left(-\sin\phi\frac{\partial}{\partial x} + \cos\phi\frac{\partial}{\partial y}\right)\left(\cos\phi\frac{\partial}{\partial x} + \sin\phi\frac{\partial}{\partial y}\right)\,. \end{aligned} \tag{14}$$

Evaluated and expressed in terms of d/dl_1 and d/dl_2, the Lie bracket becomes

$$\left[\frac{d}{dl_1},\frac{d}{dl_2}\right] = -\left(\frac{d\phi}{dl_1}\frac{d}{dl_1} + \frac{d\phi}{dl_2}\frac{d}{dl_2}\right) \neq 0\,. \tag{15}$$

This means that the two directional fields, though mutually orthogonal, have not been properly parameterized to form a coordinate system. Let

$$\frac{d}{du} = \Lambda_1\frac{d}{dl_1}\,, \qquad \frac{d}{dv} = \Lambda_2\frac{d}{dl_2} \tag{16}$$

be the coordinate system along the same characteristic directions of the image, with (u, v) the orthogonal, properly parameterized coordinates and the two functions Λ_1, Λ_2 to be determined. Recalculate their Lie bracket:

$$\begin{aligned}\left[\frac{d}{du},\frac{d}{dv}\right] &= \frac{d}{du}\frac{d}{dv} - \frac{d}{dv}\frac{d}{du} = \frac{d}{du}\left(\Lambda_2\frac{d}{dl_2}\right) - \frac{d}{dv}\left(\Lambda_1\frac{d}{dl_1}\right)\\ &= \Lambda_1\Lambda_2\left(\frac{d}{dl_1}\frac{d}{dl_2} - \frac{d}{dl_2}\frac{d}{dl_1}\right) + \frac{d\Lambda_2}{du}\frac{d}{dl_2} - \frac{d\Lambda_1}{dv}\frac{d}{dl_1}\\ &= \left(-\Lambda_2\frac{d\phi}{du} - \frac{d\Lambda_1}{dv}\right)\frac{d}{dl_1} + \left(-\Lambda_1\frac{d\phi}{dv} + \frac{d\Lambda_2}{du}\right)\frac{d}{dl_2}\,, \end{aligned} \tag{17}$$

where (15) is used in the last step. For (u, v) to be coordinates, (17) is required to be identically zero (from now on we write $\partial/\partial(\cdot)$ instead of $d/d(\cdot)$ to indicate that the directional derivatives of the u, v variables are, in addition, coordinate or "partial" derivatives):

$$\frac{\partial \phi}{\partial u} = -\frac{1}{\Lambda_2}\frac{\partial \Lambda_1}{\partial v} , \qquad \frac{\partial \phi}{\partial v} = \frac{1}{\Lambda_1}\frac{\partial \Lambda_2}{\partial u} . \tag{18}$$

This, along with (13), gives rise to the following useful identities:

$$\frac{\partial \phi}{\partial u} = -\frac{1}{\lambda_2 \Lambda_2}\frac{\partial(\lambda_1 \Lambda_1)}{\partial v} , \qquad \frac{\partial \phi}{\partial v} = \frac{1}{\lambda_1 \Lambda_1}\frac{\partial(\lambda_2 \Lambda_2)}{\partial u} ; \tag{19}$$

$$\frac{\partial(\lambda_1 \Lambda_1)}{\partial v} = \lambda_2 \frac{\partial \Lambda_1}{\partial v} , \qquad \frac{\partial(\lambda_2 \Lambda_2)}{\partial u} = \lambda_1 \frac{\partial \Lambda_2}{\partial u} ; \tag{20}$$

$$\frac{\partial \log \Lambda_1}{\partial v} = -\frac{1}{\lambda_1 - \lambda_2}\frac{\partial \lambda_1}{\partial v} , \qquad \frac{\partial \log \Lambda_2}{\partial u} = \frac{1}{\lambda_1 - \lambda_2}\frac{\partial \lambda_2}{\partial u} . \tag{21}$$

The last equation (21) specifies, at least in theory, the two unknown functions Λ_1 and Λ_2 up to a freedom to be discussed in the next section.

The integrability conditions for ϕ (i.e. the exchangeability of its mixed second derivatives) associated with (18) and (19) are as follows:

$$\frac{\partial}{\partial u}\left(\frac{1}{\Lambda_1}\frac{\partial \Lambda_2}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{1}{\Lambda_2}\frac{\partial \Lambda_1}{\partial v}\right) = 0 , \tag{22}$$

$$\frac{\partial}{\partial u}\left(\frac{1}{\lambda_1 \Lambda_1}\frac{\partial(\lambda_2 \Lambda_2)}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{1}{\lambda_2 \Lambda_2}\frac{\partial(\lambda_1 \Lambda_1)}{\partial v}\right) = 0 . \tag{23}$$

These equations may yield formal solutions under certain circumstances, but they do not impose further constraints on the two unknowns Λ_1, Λ_2.

3 Gauge Freedom in Image Coordinates

So far, we have derived, for an arbitrary image, the set of orthogonal curves $u = u_0$ and $v = v_0$ which correspond to directional flows of the image Hessian. These curves are dependent and only dependent on the image function $f(x, y)$; hence they are specific to a given image. This curvilinear coordinate net (u, v), together with the critical points of the image Hessian (either umbilic points $\lambda_1 = \lambda_2$, or degenerate points $\lambda_1 \lambda_2 = 0$), form the "signature" of a grey-level image. This characteristic flow "portrait" is called *image coordinates.*

The perceptual metric (1), being the square of the image Hessian, is diagonal under this new, image-dependent, coordinate system (u, v):

$$g_{\mu\nu} = \begin{pmatrix} (\lambda_1 \Lambda_1)^2 & \\ & (\lambda_2 \Lambda_2)^2 \end{pmatrix} , \tag{24}$$

so that the line element (under the perceptual metric) can be written as

$$ds^2 = (\lambda_1 \Lambda_1)^2 \, du^2 + (\lambda_2 \Lambda_2)^2 \, dv^2 . \tag{25}$$

On the other hand, the line element under the physical metric ($d\bar{s}^2 = dx^2 + dy^2$, denoted by an overhead bar) is now expressible using (u, v) coordinates as

$$d\bar{s}^2 = \Lambda_1^2 \, du^2 + \Lambda_2^2 \, dv^2 \ , \tag{26}$$

with the metric tensor

$$\bar{g}_{\mu\nu} = \begin{pmatrix} \Lambda_1^2 & \\ & \Lambda_2^2 \end{pmatrix} \ . \tag{27}$$

From (27) and (24), one can see that the curvilinear image coordinates (u, v) are orthogonal under both the physical and the perceptual metric (their respective metric is diagonalized). In fact, these two families of coordinating curves are unique in that they are orthogonal everywhere and in any sense (physical or perceptual). As a comparison, the image-independent coordinates (x, y), which render trivial the physical metric (identity matrix), are not best suited to express the perceptual metric (resulting in the complicated Hessian matrix). Likewise, though the geodesic coordinates $(X, Y) = (f_x, f_y)$ trivialize the perceptual metric as in [8], they are not orthogonal under the image-independent physical metric.

As good image coordinates as (u, v) are, the metric tensor (either perceptual or physical) is not trivialized, but merely diagonalized. Furthermore, they are only partially determined, which will now be discussed. Note that there is a basic freedom in specifying any coordinate system, that is, the freedom of arbitrarily (but individually) stretching the two coordinating lines. If $\partial/\partial u$, $\partial/\partial v$ are the coordinate bases, then $A(u)\,\partial/\partial u$, $B(v)\,\partial/\partial v$ are also coordinates for arbitrary functions $A(u)$ and $B(v)$; it is easy to verify that their Lie bracket vanishes identically. In the present case, this freedom is reflected in the solutions of $\Lambda_1(u, v)$ and $\Lambda_2(u, v)$ from (18). To see this, let

$$\frac{\partial}{\partial u} \to A(u)\,\frac{\partial}{\partial u} \ , \qquad \frac{\partial}{\partial v} \to B(v)\,\frac{\partial}{\partial v} \ , \tag{28}$$

then

$$\Lambda_1(u, v) \to A(u)\,\Lambda_1(u, v) \ , \qquad \Lambda_2(u, v) \to B(v)\,\Lambda_2(u, v) \tag{29}$$

also satisfy (18), the equation to re-parameterize the flow fields. The directional derivatives (5) remain unaffected (see (16)). Equations (28) and (29) together form what can be called a *gauge transformation* of the image coordinates.

To fix a particular gauge, the following approach may be adopted. The Laplacian of any function in orthogonal coordinates under metric tensor $G_{\mu\nu}$ is given by (e.g., [2, p. 41])

$$\Delta = \frac{1}{\sqrt{G_{11}G_{22}}} \left(\frac{\partial}{\partial u} \left(\sqrt{G_{11}G_{22}}\, G^{11} \frac{\partial}{\partial u} \right) + \frac{\partial}{\partial v} \left(\sqrt{G_{11}G_{22}}\, G^{22} \frac{\partial}{\partial v} \right) \right) \ . \tag{30}$$

Applying the perceptual metric (24) yields:

$$\Delta\phi(u, v) = \frac{1}{(\lambda_2\Lambda_2)(\lambda_1\Lambda_1)} \left\{ \frac{\partial}{\partial u} \left(\frac{\lambda_2\Lambda_2}{\lambda_1\Lambda_1} \frac{\partial\phi}{\partial u} \right) + \frac{\partial}{\partial v} \left(\frac{\lambda_1\Lambda_1}{\lambda_2\Lambda_2} \frac{\partial\phi}{\partial v} \right) \right\}$$

$$= \frac{1}{(\lambda_2 \Lambda_2)(\lambda_1 \Lambda_1)} \left\{ \frac{\partial}{\partial u} \left(-\frac{1}{\lambda_1 \Lambda_1} \frac{\partial(\lambda_1 \Lambda_1)}{\partial v} \right) + \frac{\partial}{\partial v} \left(\frac{1}{\lambda_2 \Lambda_2} \frac{\partial(\lambda_2 \Lambda_2)}{\partial v} \right) \right\}$$

$$= \frac{1}{(\lambda_2 \Lambda_2)(\lambda_1 \Lambda_1)} \frac{\partial^2}{\partial u \partial v} \left(\log \frac{\lambda_2 \Lambda_2}{\lambda_1 \Lambda_1} \right) . \tag{31}$$

It can easily be verified that $\Delta\phi$ is a gauge-invariant quantity. It is suggested that a class of image functions ("good" or Gestalt images) exist such that

$$\Delta\phi = 0 . \tag{32}$$

This makes

$$\frac{\partial^2}{\partial u \partial v} \left(\log \frac{\lambda_2 \Lambda_2}{\lambda_1 \Lambda_1} \right) = 0 , \tag{33}$$

or

$$\log \left(\frac{\lambda_2 \Lambda_2}{\lambda_1 \Lambda_1} \right) = \log A(u) - \log B(v) . \tag{34}$$

The arbitrary functions $A(u)$, $B(v)$ may always be absorbed into Λ_1, Λ_2 respectively (due to the basic gauge freedom), giving rise to the perceptual gauge for Gestalt images:

$$\lambda_1 \Lambda_1 = \lambda_2 \Lambda_2 \; (= \Omega) . \tag{35}$$

From (19), it can been seen that $\phi(u, v)$ and $\log \Omega(u, v)$ satisfy Cauchy-Riemann equations:

$$\frac{\partial \phi}{\partial u} = -\frac{\log \Omega}{\partial v} , \qquad \frac{\partial \phi}{\partial v} = \frac{\log \Omega}{\partial u} . \tag{36}$$

The line element of the perceptual metric now becomes

$$ds^2 = \Omega^2 \left(du^2 + dv^2 \right) . \tag{37}$$

This is to say that, when the perceptual gauge (35) is satisfied, the perceptual metric is merely an isothermal mapping under image coordinates (u, v).

To find the relation between image coordinates (u, v) and geodesic coordinates $(X, Y) = (f_x, f_y)$ (as in [8]), apply the chain rule of differentiation and observe:

$$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)^T = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} \left(\frac{\partial}{\partial X}, \frac{\partial}{\partial Y} \right)^T . \tag{38}$$

An application of (3), (5), and (16) yields

$$\left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v} \right)^T = \begin{pmatrix} \lambda_1 \Lambda_1 & \\ & \lambda_2 \Lambda_2 \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \left(\frac{\partial}{\partial X}, \frac{\partial}{\partial Y} \right)^T . \tag{39}$$

Under the perceptual gauge (35), equation (39) can be recast after introducing complex variables $w = u + iv$, $Z = X + iY$:

$$dw = (\Omega \exp(i\phi))^{-1} dZ = \exp(-(\log \Omega + i\phi))\, dZ . \tag{40}$$

Since $\log \Omega$ and ϕ form a Cauchy-Riemann pair (see (36)), the above expression indicates that the image coordinates (u, v) and the geodesic coordinates (X, Y) are conformally related.

A similar derivation may be carried out using the physical metric $\bar{g}_{\mu\nu}$. The analogous equation for (31) is

$$\bar{\Delta}\phi(u,v) = \frac{1}{\Lambda_1\Lambda_2}\frac{\partial^2}{\partial u\partial v}\left(\log\frac{\Lambda_2}{\Lambda_1}\right) , \tag{41}$$

which leads to the following physical gauge:

$$\Lambda_1 = \Lambda_2 \ (= \Lambda) \tag{42}$$

and the corresponding Cauchy-Riemann pair:

$$\frac{\partial\phi}{\partial u} = -\frac{\log\Lambda}{\partial v} , \qquad \frac{\partial\phi}{\partial v} = \frac{\log\Lambda}{\partial u} . \tag{43}$$

Analogous to (40), the new coordinates (u, v) are now conformally related to the Cartesian coordinates (x, y) through (after introducing $z = x + iy$):

$$dw = (\Lambda\exp(i\phi))^{-1}dz = \exp(-(\log\Lambda + i\phi))\,dz . \tag{44}$$

The conformal relationships (40) under the perceptual gauge or (44) under the physical gauge indicate that ϕ — and hence (u, v) — will be completely specified given the boundaries and critical points of the mapping. These include locations where the image Hessian is umbilic ($\lambda_1 = \lambda_2$) or degenerate ($\lambda_1\lambda_2 = 0$).

4 Manipulations Using Image Coordinates

The curvilinear image coordinates (u, v) involve only the second derivatives with respect to space. Hence, they are covariant under an affine transformation (or mapping) of the two-dimensional visual space. They are "centred" on the visual image — in fact, the two sets of curves $u = u_0$ or $v = v_0$ represent "line-drawings" of a grey-level image. We now intend to relate their curvature, a geometrical descriptor, to the original image function. The intrinsic (geodesic) curvatures of the orthogonal coordinating curves $u = u_0$ and $v = v_0$ are (e.g., [7, p.130]):

$$k_1 = -\frac{1}{\sqrt{G_{11}G_{22}}}\frac{\partial\sqrt{G_{11}}}{\partial v} \quad \text{along } u = u_0 , \tag{45}$$

$$k_2 = \frac{1}{\sqrt{G_{11}G_{22}}}\frac{\partial\sqrt{G_{22}}}{\partial u} \quad \text{along } v = v_0 . \tag{46}$$

Under the perceptual metric $g_{\mu\nu}$, they are

$$k_1 = \frac{1}{\lambda_1\Lambda_1}\frac{\partial\phi}{\partial u} , \qquad k_2 = \frac{1}{\lambda_2\Lambda_2}\frac{\partial\phi}{\partial v} . \tag{47}$$

Likewise, under the physical metric $\bar{g}_{\mu\nu}$,

$$\bar{k}_1 = \frac{1}{\Lambda_1}\frac{\partial\phi}{\partial u} , \qquad \bar{k}_2 = \frac{1}{\Lambda_2}\frac{\partial\phi}{\partial v} . \tag{48}$$

Therefore

$$\bar{k}_1 = \lambda_1 k_1 \ , \qquad \bar{k}_2 = \lambda_2 k_2 \ . \tag{49}$$

Topological mappings of the visual manifold that preserve the neighbourhood relationship, including translation, rotation, scaling, and rubber-sheet deformation, can now be described as Lie dragging of the image "line-drawings". In particular, the two vector fields $A(u)\partial/\partial u$ and $B(v)\partial/\partial v$ themselves form the commutative bases for Lie transformation groups operating on an image or its associated sensory fields. They capture local symmetry of an image and express all sensible rubber-sheet deformations (see Sect. 1).

The image coordinates can be used to compute the intrinsic (Gaussian) curvature of a space via the well-known Gauss equation (e.g., [7, p.113])

$$K = -\frac{1}{2\sqrt{G_{11}G_{22}}} \left\{ \frac{\partial}{\partial v} \left(\frac{1}{\sqrt{G_{11}G_{22}}} \frac{\partial G_{11}}{\partial v} \right) + \frac{\partial}{\partial u} \left(\frac{1}{\sqrt{G_{11}G_{22}}} \frac{\partial G_{22}}{\partial u} \right) \right\} \ . \tag{50}$$

For the perceptual metric (24), a straightforward calculation using (23) yields

$$K = 0 \ . \tag{51}$$

This is consistent with the fact that the Riemann curvature tensor of the perceptual space vanishes identically [8]. There is no intrinsic curviness for the perceptual space. Since the Gaussian curvature is an intrinsic geometrical quantity, any three-dimensional embedding of the two-dimensional perceptual geometry, if possible, would have to be a developable surface. By the way, a similar calculation demonstrates, quite trivially, that the Gaussian curvature $\bar{K}$ under the physical metric (27) is also identically zero.

5 Patterns with Circular Symmetry: an Example

As a concrete example of this approach, let us calculate the image coordinates of a circularly symmetric pattern $f(x, y) = F(r)$, with $r \equiv \sqrt{x^2 + y^2}$. The second derivatives (image Hessian) are

$$f_{xx} = \frac{x^2}{r^2}F'' + \frac{y^2}{r^3}F' \ , \quad f_{xy} = \frac{xy}{r^2}F'' - \frac{xy}{r^3}F' \ , \quad f_{yy} = \frac{y^2}{r^2}F'' + \frac{x^2}{r^3}F' \ , \tag{52}$$

with eigenvalues easily found to be

$$\lambda_1 = F''(r) \ , \qquad \lambda_2 = F'(r)/r \ . \tag{53}$$

Their corresponding eigenvectors simply point along the radial and angular directions respectively,

$$\mathbf{n}_1 = [x/r, y/r]^T \ , \qquad \mathbf{n}_2 = [y/r, -x/r]^T \ . \tag{54}$$

Denoting $\theta \equiv \arctan(y/x)$, the two directional derivatives in (5) become

$$\frac{d}{dl_1} = \frac{\partial}{\partial r} \ , \qquad \frac{d}{dl_2} = \frac{1}{r}\frac{\partial}{\partial \theta} \ . \tag{55}$$

Equation (13) can be verified as being satisfied. To solve for the unknown, parameterizing functions Λ_1, Λ_2, apply (16), (55) and (21):

$$\frac{\partial \log \Lambda_1}{\partial \theta} = -\frac{1}{\lambda_1 - \lambda_2}\frac{\partial \lambda_1}{\partial \theta} = 0 , \tag{56}$$

$$\frac{\partial \log \Lambda_2}{\partial r} = \frac{1}{\lambda_1 - \lambda_2}\frac{\partial \lambda_2}{\partial r} = \frac{1}{r} . \tag{57}$$

The solutions are, along with arbitrary functions $A(r)$ and $B(\theta)$,

$$\Lambda_1 = A(r) , \qquad \Lambda_2 = rB(\theta) , \tag{58}$$

so that

$$\frac{\partial}{\partial u} = A(r)\frac{\partial}{\partial r} , \qquad \frac{\partial}{\partial v} = B(\theta)\frac{\partial}{\partial \theta} . \tag{59}$$

Therefore, the variables r and θ (or arbitrary functions of either) are indeed image coordinates for circularly symmetric images, though they were previously introduced merely as shorthand notations of given functions of (x, y). In this case, both $\Delta\phi$ and $\bar{\Delta}\phi$ equal zero (see (33) and (41)); thus patterns of circular symmetry are good "Gestalt" figures. It can be shown that $(u, v) = (\log F', \theta)$ and $(u, v) = (\log r, \theta)$ under perceptual and physical gauges respectively.

References

1. Ferraro, M., Caelli, T.M. (1988). Relationship between integral transform invariances and Lie group theory, J. Opt. Soc. Am. A 5, pp. 738–742.
2. Göckeler, M., Schücker T. (1987). Differential Geometry, Gauge Theories, and Gravity, Cambridge University Press, Cambridge, UK.
3. Hoffman, W.C. (1966). The Lie algebra of visual perception, J. Math. Psychol. 3, pp. 65–98.
4. Pintsov, D.A. (1989). Invariant pattern recognition, symmetry, and Radon transforms, J. Opt. Soc. Am. A 6, pp. 1544–1554.
5. Schutz, B.F. (1980). Geometrical Methods of Mathematical Physics, Cambridge University Press, Cambridge, UK.
6. Segman, J. (1992). Fourier cross correlation and invariance transformations for an optimal recognition of functions deformed by affine groups, J. Opt. Soc. Am. A 9, pp. 895–902.
7. Struik, D.J. (1950). Lectures on Classical Differential Geometry, Addison-Wesley, Massachusettes. Republished in 1988 by Dover, New York.
8. Zhang, J., Wu, S. (1990). Structure of visual perception, Proc. Natl. Acad. Sci. USA. 87, pp. 7819–7823.