Journal of Mathematical Imaging and Vision, 3, 327–348 (1993).

Cartesian Differential Invariants in Scale-Space

L.M.J. FLORACK, B.M. TER HAAR ROMENY, J.J. KOENDERINK, AND M.A. VIERGEVER
University of Utrecht, Utrecht, The Netherlands

Abstract. We present a formalism for studying local image structure in a systematic, coordinate-independent, and robust way, based on scale-space theory, tensor calculus, and the theory of invariants. We concentrate on *differential invariants*. The formalism is of general applicability to the analysis of grey-tone images of various modalities, defined on a D-dimensional spatial domain.

We propose a "diagrammar" of differential invariants and tensors, i.e., a diagrammatic representation of image derivatives in scale-space together with a set of simple rules for representing meaningful local image properties. All local image properties on a given level of inner scale can be represented in terms of such diagrams, and, vice versa, all diagrams represent coordinate-independent combinations of image derivatives, i.e., true image properties.

We present *complete* and *irreducible sets* of (nonpolynomial) differential invariants appropriate for the description of local image structure up to any desired order. Any differential invariant can be expressed in terms of *polynomial invariants*, pictorially represented by closed diagrams. Here we consider a complete, irreducible set of polynomial invariants up to second order (inclusive).

Examples of differential invariants up to fourth order (inclusive), calculated for synthetic, noise-perturbed, 2-dimensional test images, are included to illustrate the main theory.

Key words. Cartesian differential invariants, Cartesian tensors, irreducible invariants, Gaussian scale-space, local image structure

1 Introduction

In image analysis a distinction can often be made between *local* and *multilocal* methods. The key idea in this dichotomy is *inner scale:* local image properties can be associated with a *single base point* at a *given inner scale* (i.e., inverse resolution), comprising a *local neighbourhood*. Such properties can be defined in terms of spatial derivatives taken at a given base point,[1] and so in our context "local" does not mean "punctal." To extract a derivative one needs an integration filter defined on a full spatial neighbourhood of the base point of interest, with an effective extent proportional to the inner scale. Multilocal properties are associated with multiple local neighbourhoods.

The intrinsic scale degree of freedom accounts for ambiguities in this interpretation. For example, the average image intensity appears to be a multilocal property on a small scale, but it turns out to have a local interpretation on a sufficiently large scale. The principle of digital halftoning or dithering used for rendering the illusion of grey-tone pictures by using only black and white picture elements, clearly demonstrates the virtue of this ambiguity.

Although many image properties do not have a local interpretation, a thorough understanding of local image properties is of fundamental importance. A rigorous approach towards understanding local image issues also permits a connection between local image analysis and the well-established mathematical literature on differential methods. These methods can be applied to local image analysis in a straightforward way. Without claiming that it is complete in any way, we list some references that are relevant in this context: [1]–[3] (differential geometry), [4]–[13] (algebraic invariants), [14]–[19] (discriminants and resultants), [20] (binary forms), [21] (irreducible invariants), [22], [23] (the theory of invariants in general).

Our goal is to provide basic information needed for appreciation and operationalisation of the well-established differential methods, applied to front-end vision and local image analysis. To this end we need to (i) *operationalise differential methods according to some well-posed scheme* and (ii) *explain the principle of invariance underlying a coordinate-independent approach.* It is crucial for any local image operation to be driven by operators that are well posed by construction and to operate in a geometrically meaningful way.

At some higher level of description one typically uses *a priori* assumptions about the grey-level image, e.g. the knowledge that the image is actually obtained through a projection or that lighting conditions affect the grey-level structure but not the essence of the underlying scene. In some cases one may want to disregard the influence of monotonic grey-level transformations to reveal the image's structure independent of gamma corrections or at least to ignore the effect of a constant rescaling and/or offset of grey values. These considerations generally lead to a more restrictive class of so-called *invariants*, i.e., image properties that are insensitive to variations of the irrelevant parameters.

The way in which this comes about in the analysis is by postulating a *group* that transforms the parameters one would like to disregard as being irrelevant for the interpretation and by requiring invariance under the action of that group. Some examples are the (plane) affine and projective groups often considered in contexts of projection data [24], the "stereo groups" GS (4) and SS (4) introduced in [25], and the group of general intensity transformations [26].

The introduction of such groups implies a redundancy of image data: not all of the image's grey-level structure is considered to be relevant. *In this paper we disregard any such a priori information about the image and instead look for a complete representation of its local grey-level structure.* Such a basic syntactical representation naturally precedes any semantic level of description. Consequently, we require invariance only under the group of Cartesian transformations, i.e., SO $(D) \times \mathrm{T}(D)$, in which $\mathrm{SO}(D)$ denotes the *special orthogonal group* and $\mathrm{T}(D)$ the *translation group* in D spatial dimensions.

Of course, we can always set up a Cartesian coordinate frame at each base point of the image in order to carry out any local calculations, but we have to make sure that the results of these calculations are actually independent of this particular frame (for only then do they correspond to intrinsic local image properties). Requiring invariance under the group of Cartesian coordinate transformations will guarantee this. This underlines the general importance of the Cartesian gauge group and its invariant objects (*Cartesian tensors*) in a front-end vision processor. In particular, all local image descriptors in a grey-scale image, be it affine, projective, or whatever invariants, can be expressed in terms of Cartesian tensors.

Spatial scale is rigorously defined by means of a construct known as *scale-space*, which is generated by a Gaussian kernel [27]–[30], the width of which corresponds to the inner scale. Using a normalised Gaussian as the scale-space generating kernel uniquely guarantees that no spurious detail is created in the fine-to-coarse direction. Alternatively, a normalised Gaussian is the unique smooth kernel that is *self-similar* along the resolution dimension, which accounts for the *a priori* equivalence of all scales.

A complete family of local neighbourhood operators can be constructed by a local method known as the "prolongation" [31] of the basic scale-space kernel. In principle, prolongation can be carried out up to any order N and yields a set of Gaussian derivative filters that implement the notion of a *local jet* of order N [32]. In this local context the Gaussian appears merely as the (unique) zeroth-order member of a complete family of local neighbourhood operators or *scaled differential operators* known as the *Gaussian family* [33], [34]. The higher-order local neighbourhood operators are linear derivatives of this basic one.

The organisation of the paper is as follows. We first review the Gaussian family and discuss the concept of a local jet in section 2. Here, we also discuss the matters of well-posedness and completeness. Then, in section 3 we introduce the notions of Cartesian tensors and invariants, as well as a convenient diagrammatic representation of these. The main goal here is to illustrate the role of the tensor formalism in local image analysis. A more formal treatment of tensor calculus is given in appendix A. Special attention is paid to irreducible invariants. The proof of one proposition on this is somewhat more involved and is postponed to appendix B. Finally, section 4 contains some examples of invariants applied to synthetic images.

2 Basic Concepts

2.1 The Gaussian Family and Its Cartesian Representation

We need a more precise definition of what we mean by a "local image property" or, equivalently, what we mean by a "local neighbourhood operator," used to extract—and thereby define—a local image property. To capture *all* local image properties we need a *complete* family of such local neighbourhood operators.

In this paper we use the following *Cartesian* representation of the Gaussian family (see [33] for some other useful coordinate representations):

DEFINITION 1 (Cartesian family). A Complete, hierarchically ordered family of multiplicative scale-space kernels in D dimensions is given in the Fourier representation by the set

$$\{\tilde{G}_{i_1\cdots i_n}(\omega;\sigma) = i\omega_{i_1}\cdots i\omega_{i_n}\tilde{G}(\omega;\sigma) \mid (\omega;\sigma) \in \mathbf{R}^D \times \mathbf{R}^+,\, n \in \mathbf{Z}_0^+\},$$

in which the zeroth-order member is given by

$$\tilde{G}(\omega;\sigma) = \exp\left(-\frac{1}{2}\sigma^2\omega^2\right).$$

Alternatively, in the spatial representation it is given by the following set of convolution filters:

$$\{G_{i_1\cdots i_n}(\mathbf{x};\sigma) = \partial_{i_1\cdots i_n}G(\mathbf{x};\sigma) \mid (\mathbf{x};\sigma) \in \mathbf{R}^D \times \mathbf{R}^+,\, n \in \mathbf{Z}_0^+\},$$

with

$$G(x;\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}^D}\exp\left(-\frac{1}{2}\frac{x^2}{\sigma^2}\right).$$

In this definition and henceforth we use the following notation for the spatial coordinates and frequencies, respectively: $\mathbf{x} = (x_1,\ldots,x_D), \omega = (\omega_1,\ldots,\omega_D)$. Also, we use Latin indices from the middle of the alphabet, each of which can take values in the range $1,\ldots,D$, to label a corresponding Cartesian coordinate. For this reason we speak of *Cartesian indices*. Inner scale is parametrised by the Gaussian width parameter σ. By $\partial_{i_1\cdots i_n}$ we mean the linear partial derivative operator $\partial^n/\partial x_{i_1}\cdots\partial x_{i_n}$, etc. Throughout the paper we make use of the Einstein summation convention with respect to these indices: if an index occurs twice in a given term, a summation over all possible index values is tacitly assumed, so that $Q_{ii} \stackrel{\text{def}}{=} \sum_{i=1}^{D} Q_{ii}$.

The index n imposes a (representation-independent) hierarchy on the Gaussian family and corresponds to the order of differentiation in the spatial domain. However, there is no natural hierarchy among the kernels within a fixed order n, and the multiplicity of this degeneracy, i.e., the number of independent, so-called *essential components* among the fixed order kernels $G_{i_1\cdots i_n}$, depends on the order and on the dimension of space and is given by the following theorem.

THEOREM 1 (essential components of nth-order kernel). The number of essential components of the nth-order kernel $G_{i_1\cdots i_n}$ in D spatial dimensions and on a given level of scale is given by

$$\overset{(D)}{m}_n = \binom{n+D-1}{D-1}.$$

This follows from the fact that the kernels are symmetric with respect to the interchange of any pair of indices. Theorem 1 gives the number of kernels minimally required (and also sufficient) for a full nth-order description of local image structure on a given level of scale.

The Gaussian derivative filters functionally correspond to "blown-up" or *scaled differential operators* in a precise sense, viz., they are obtained by an (isotropic) diffusion of the conventional differential operators, the evolution parameter s of which corresponds to the inner scale σ according to $2s = \sigma^2$.

The conventional operators can thus be thought of as the hypothetical zero-scale initial conditions to the scale-space generating diffusion equation

$$\begin{cases} (\Delta - \partial_s)\phi = 0, \\ \phi(\mathbf{x}; s = 0) = \partial_{i_1 \cdots i_n}\delta(\mathbf{x}), \end{cases} \quad (1)$$

in which $\delta(\mathbf{x})$ is the well-known *Dirac distribution*. The solution to this initial-value problem precisely corresponds to the Cartesian family of Definition 1: $\phi(\mathbf{x}; s) = G_{i_1 \cdots i_n}(\mathbf{x}; \sigma = \sqrt{2s})$.

2.2 *The Local Jet and its Cartesian Representation*

A local jet of order N of a given function $f \in C^N$ (Dom f), denoted by $J^N[f]$, is defined as the equivalence class of functions that have the same local structure up to order N (N inclusive). In other words, all images in a given local N-jet (and on a given level of scale) are locally indistinguishable when we have only the Gaussian derivatives of orders $0, \ldots, N$ at our disposal for the extraction of local information. The operational definition of a local jet for images is given by the following.

DEFINITION 2 (local jet of order N). Let $\psi : \mathbf{R}^D \to \mathbf{R})$ be a given image, and let σ be a physically sensible inner scale for ψ. Then the local jet of ψ of order N at base point $\mathbf{x}$ and inner scale σ, denoted by $J^N[\psi](\mathbf{x}; \sigma)$, can be represented with respect to an arbitrary Cartesian coordinate system by the set

$$J^N[\psi](\mathbf{x}; \sigma) = \{L_{i_1 \cdots i_n}(\mathbf{x}; \sigma) | (\mathbf{x}; \sigma) \in \mathbf{R}^D \times \mathbf{R}^+, n = 0, \ldots, N\},$$

in which $L_{i_1 \cdots i_n}$ is given by the convolution of ψ with the Gaussian derivative $G_{i_1 \cdots i_n}(.; \sigma)$:

$$L_{i_1 \cdots i_n}(\mathbf{x}; \sigma) = (G_{i_1 \cdots i_n} * \psi)(\mathbf{x}; \sigma).$$

Alternatively, the local jet of order N can be represented in Fourier space by the set

$$\mathcal{J}^N[\tilde{\psi}](\omega; \sigma) = \{\tilde{L}_{i_1 \cdots i_n}(\omega; \sigma) | (\omega; \sigma) \in \mathbf{R}^D \times \mathbf{R}^+, n = 0, \ldots, N\},$$

in which $\tilde{L}_{i_1 \cdots i_n}$ is given by the product of $\tilde{\psi}$, i.e., the Fourier transform of ψ, with the Gaussian derivative in Fourier space, $\tilde{G}_{i_1 \cdots i_n}(.; \sigma) = i\omega_{i_1} \cdots i\omega_{i_n} \tilde{G}(.; \sigma)$:

$$\tilde{L}_{i_1 \cdots i_n}(\omega; \sigma) = \left(\tilde{G}_{i_1 \cdots i_n} \tilde{\psi}\right)(\omega; \sigma).$$

Definition 2 seems somewhat sloppy since it seems to rely on the choice of Cartesian coordinates. The local jet is really a coordinate-independent object, however, and we show in section 3 how to promote the index notation to a symbolic, coordinate-independent status by associating a Cartesian transformation with each index, so that there is no need for a modification. For the moment, however, we ignore this and discuss a number of aspects that readily follow from this definition.

The meaning of a "physically sensible" inner scale in Definition 2 is that σ should be fairly within the physical limits set by the imaging modality's resolving power (the grid or pixel size or noise correlation width ε) and the typical size R of its field of view. Although we formally have $\lim_{\sigma \downarrow 0} L(\mathbf{x}; \sigma) = \psi(\mathbf{x})$, the accuracy of any physical representation of $L(\mathbf{x}; \sigma)$ is limited to $\mathcal{O}(\varepsilon/\sigma)$ at the very best, and so this zero-scale limit never makes any real sense in practice (but, of course, it is the only sensible *a priori* requirement if we do not want to rely on device dependencies). In all that follows we implicitly assume that σ satisfies $\max(\varepsilon/\sigma, \sigma/R) \ll 1$ and ignore all device-specific effects of these orders (cf. [29] for the small-scale boundary layer). Also, since we are primarily interested in invariance principles, we leave open the question of how to constrain N in Definition 2 when we are given spatial and intensity resolution limitations. Again, any value is conceivable under appropriate imaging conditions. High-order derivatives are generally feasible and robust, provided that they can be calculated on a sufficiently high scale (relative to pixel scale and noise correlation width) and provided that we have a sufficient resolution of intensity values (dynamic resolution, noise). We refer to [35] for a detailed discussion of these trade-offs.

From Theorem 1 it follows that the total number of local degrees of freedom per scale in the N-jet equals $\overset{(D)}{\mu}_N = \sum_{n=0}^{N} \overset{(D)}{m}_n$. Note that we could have written $L_{i_1 \cdots i_n}$ as $\partial_{i_1 \cdots i_n}\{G * \psi\}$ or as $G * \psi_{i_1 \cdots i_n}$, i.e., as the derivative of the blurred image $L = G * \psi$ and as the blurred derivative image, respectively. This reveals notational consistency and clearly shows the functionality of the Gaussian family as an apparatus of differential calculus for physical signals in the exact sense [36].

2.3 Well-Posed Differentiation

It is well-known that differentiation in the conventional sense, by its very definition, is ill posed in the sense of Hadamard. This means that an insignificant perturbation of input data may have an arbitrarily large effect on the output. As an example, consider two, almost equal smooth functions (with respect to a suitable norm) $f_1(x) = f(x)$ and $f_2(x) = f(x) + \delta f(x)$, where $\delta f(x)$ is a small, additive, high-frequency perturbation, say, $\delta f(x) = \varepsilon g(x/\varepsilon^{a+1})$, with $0 < \varepsilon \ll 1$ and $0 < \max(|g(x)|, |g'(x)|) < M$ for all x. So $\delta f(x)$ can be made insignificantly small by taking a suitably small ε and can even be made to vanish completely by taking the limit $\varepsilon \downarrow 0$. However, if $a > 0$, then $\delta f'(x) = g'(x/\varepsilon^{a+1})/\varepsilon^a$ can become arbitrarily large and generally does not even exist for $\varepsilon \downarrow 0$. This shows that classical differentiation does not have the slightest flexibility with respect to small variations of the input data, which makes it useless for image analysis. Given the intrinsically finite resolution of any physical observation, it clearly makes no sense to use ill-posed operators that are extremely sensitive to infinitesimal input variations.

It is not commonly appreciated that the restriction to smooth functions does not help because ill-posedness has nothing to do with the choice of a function space (i.e., the operands); it is an artefact of the *operations*. Therefore we need a modification of the differential operators and *not a smoothing or regularisation of the image*. In our opinion, there is no justification for modifying the image data before the analysis (we will not question the image acquisition and reconstruction stages). When operationalised in the sense of Definition 2, the process of differentiation clearly becomes well posed. The inevitable price we have to pay for this is the extra scale degree of freedom. But this freedom truly makes sense if we recall the example of a spatially dithered halftone image, which appears smooth provided that we keep sufficient distance so as to be unable to resolve the binary picture elements. From the apparent smoothness of such an image when viewed from a distance, one should be inclined to take derivatives. The procedure for doing this is thus in some sense opposite to the usual one in mathematics. We have to resort to a sufficiently large inner scale instead of zooming in on an infinitesimal or zero-scale neighbourhood: local image structure is defined by virtue of a finite aperture of intrinsically variable extent.

This physical solution to the ill-posedness problem is functionally the same as in a mathematical construct, based on so-called *tempered distributions*, which are defined as functionals on a space $\mathcal{S}$ of smooth test functions (so-called Schwartz functions [37]). One could say that physical considerations severely constrain the admissible Schwartz functions to a two-parameter family $G(\cdot, \mathbf{y}; \sigma)$ obtained by shifting (to base point $\mathbf{y}$) and dilating (by scale factor σ) a unique, smooth scaling function, the isotropic Gaussian [38]: $\Gamma(\mathbf{x}) = \left(1/\sqrt{2\pi}^D\right)e^{-x^2/2} \rightarrow \sigma^{-D}\Gamma(\sigma^{-1}(\mathbf{x}-\mathbf{y})) = G(\mathbf{x}, \mathbf{y}; \sigma)$. The shifts and dilations account for a translation and scale-invariant distribution of local operators over the entire image domain.

2.4 Completeness

Completeness of the Gaussian family is reflected in the fact that, at least in principle, the formal limit $\lim_{N\to\infty} J^N[\psi](\mathbf{x}; \sigma)$ for a given local neighbourhood $(\mathbf{x}; \sigma)$ contains all the information needed for a reconstruction of the image in a full neighbourhood of $(\mathbf{x}; \sigma)$ (by convergence of the local Taylor expansion). The attribute completeness thus relates to a *single base point* and a *fixed inner scale*.

Completeness in this formal sense is not a physically relevant issue for any particular image (there is no image for which the entire Gaussian family makes sense) but is again a necessary *a priori* requirement since there is no *a priori* restriction on the highest physically sensible order N (given any N, we can always think of an image for which the local N-jet makes accurate sense).

A physically meaningful relaxation of completeness in the strict sense is obtained by truncating at some small enough, physically sensible jet order N. Then, *by definition*, the N-truncated Gaussian family is complete on the *equivalence class* defined by the local N-jet. The necessity of truncation, induced by resolution limitations, implies that various local results ultimately have to be patched together into a global framework. This, however, is beyond the scope of the present paper. We refer to [36] for a qualitative insight into the complex problems that arise in the operationalisation of a global scheme.

3 Cartesian Invariance

3.1 Cartesian Tensors

Recall that the indices we used in Definitions 1 and 2 refer to some Cartesian coordinate system. However, both the Gaussian family and the local jet bundle should really be regarded as being coordinate-independent objects. It turns out that there is no need to modify these definitions if we interpret all indexed quantities $Q_{i_1 \ldots i_n}$ as *tensor components*, i.e., components with respect to some Cartesian coordinate system of a truly Cartesian-invariant quantity Q. It is then silently assumed that an index does not relate to a specific realisation of Cartesian axes, but rather to a representation of the *transformation group* connecting all possible realisations, i.e., the Cartesian group. This group index convention is quite common in physics and can easily be generalised to other (non-Cartesian) coordinate transformations. If the transformation group is not a subgroup of the Cartesian group, however, then this generalisation forces us to distinguish between so-called *covariant* and *contravariant tensors*, the components of which are conventionally labeled by lower (covariant) and upper (contravariant) indices, respectively.

The reason that we do not need to make the distinction between covariant and contravariant objects for the Cartesian group is the orthogonality property [see appendix A, formula (5)]. To appreciate this co- versus contra-parlance, consider what happens to the image gradient L_i on a linear transformation of a displacement vector $\tilde{x}_i = a_{ij} x_j$ (so $x_i = a_{ij}^{-1} \tilde{x}_j$). By the chain rule we have $\tilde{\partial}_i L = a_{ji}^{-1} L_j$. For a general group, this shows that x_i and L_i transform differently: x_i is said to be a contravariant vector or contravector (or simply vector), whereas L_i is called a covariant vector or covector. To avoid confusion one then introduces covariant or lower and contravariant or upper indices, and so, e.g., x_i becomes x^i. As explained in section 1, there is no strict need for considering frames other than Cartesian frames if we restrict ourselves to a description of local features. Indeed, we observe that, by virtue of the orthogonality relation $a_{ji}^{-1} = a_{ij}$, covariance and contravariance boil down to the same thing in the Cartesian case (note that the subgroup of translations has no effect on tensor indices). Henceforth we will therefore use lower indices only. For readers who are not familiar with Cartesian tensor calculus or with our notational conventions, we have added an appendix in which we have summarised the basic definitions and results of this formalism; see appendix A.

The purpose of tensor calculus is the coordinate independence of identities in group-index notation (see Claim 6 of appendix A). This is the reason that we speak of *manifest invariance* when using tensor identities. The indices used to denote the tensor components with respect to some Cartesian coordinate frame are usually symbolic in facilitating notational matters concerning tensor operations, such as contractions and multiplications. Although in principle we could manage perfectly well without the use of these indices (at the expense of having to raise the level of abstraction), the index notation has the additional benefit that it can be easily transferred into any particular coordinate system. The advantage of this is twofold. Firstly, it is of *practical convenience* if we want to perform actual calculations on images, in which case we will have to single out a particular coordinate system (e.g., one aligned with the grid in the case of a rastered image). Secondly, it is of *theoretical convenience* to be able to impose admissible constraints (i.e., constraints realisable through a Cartesian transformation) on the choice of axes in a manifest invariant way, without actually having to go through cumbersome transformations for establishing that particular coordinate system explicitly, starting from a given one; only the admissibility must be checked. Such an admissible local coordinate choice is called a *gauge choice*, and the resulting coordinates are called *gauge coordinates*. We will see later on how this gauge fixing works and how it can greatly simplify the derivation of various geometric identities.

3.2 Cartesian Invariants

We now turn to the question of how to construct (Cartesian) *invariants*. These are scalar quantities (0-tensors), associated with a single base point in the image and a given inner scale, that have coordinate-independent values. Because they are purely image induced, they constitute true local image descriptors.

The elementary building blocks for constructing these invariants for a given image are provided by the local N-jet. Assuming we have obtained the N-jet components $L_{i_1 \cdots i_p}$ for all $p = 0, \ldots, N$ with respect to some Cartesian basis, we now discuss the rules for joining these building blocks so as to obtain invariants.

Within the group-index formalism this is actually very easy. Simply form any tensor product by multiplying tensor members of the N-jet, supplemented with the two constant tensors δ_{ij} and $\varepsilon_{i_1 \cdots i_D}$, the *Kronecker* and the *Lévi–Civita tensor*, respectively (see Appendix A). Then perform a full contraction on its free indices (this, of course, requires an even number of free indices). The result is a *polynomial invariant* (or *pseudo invariant*,[2] if the product contains an odd number of Lévi–Civita tensors). The proof of this is straightforward: if $T_{i_1 \cdots i_{2n}}$ represents a $2n$-tensor obtained in the above way (possibly after a rearrangement of indices), then $T_{i_1 i_1 \cdots i_n i_n}$ transforms as $\widetilde{T}_{i_1 i_1 \cdots i_n i_n} = (\Delta) a_{i_1 j_1} a_{i_1 k_1} \cdots a_{i_n j_n} a_{i_n k_n} T_{j_1 k_1 \cdots j_n k_n}$ (in which the factor Δ, representing the determinant of a_{ij}, i.e., ± 1, shows up when dealing with a pseudo tensor). But each product $a_{i_m j_m} a_{i_m k_m} (m = 1, \ldots, n)$ in this expression is just $\delta_{j_m k_m}$ by the orthogonality property of a_{ij}, and so the right-hand side simplifies to $(\Delta) T_{i_1 i_1 \cdots i_n i_n}$.

This recipe may give us all polynomial (pseudo) invariants. Arbitrary invariants can now be obtained as functions of these polynomial ones. Of course, the number of polynomial N-jet invariants is (countably) infinite, even though we started out with only a finite number of degrees of freedom. Hence it is obvious that almost all of them are functions of a finite number of basic invariants. This explains why the index notation for a given invariant is generally not unique.

As an example of this, consider the determinant $\Delta = \frac{1}{2} \varepsilon_{ij} \varepsilon_{kl} L_{ik} L_{jl}$ of the Hessian matrix L_{ij} in two dimensions. It is well known from the calculus of matrices that we can express this in terms of the traces of L_{ij} and its square $L_{ik} L_{kj}$. This is a result of the fact that a product of two ε-tensors can be written in terms of δ-tensors only: $\varepsilon_{ij} \varepsilon_{kl} = \delta_{ik} \delta_{jl} - \delta_{il} \delta_{jk}$ (similar relations hold in higher dimensions; see appendix A). So we have $\varepsilon_{ij} \varepsilon_{kl} L_{ik} L_{jl} = L_{ii} L_{jj} - L_{ij} L_{ji}$. The dependence between the ε-tensor and the δ-tensor is a clear cause of the notational plurality. A less obvious example is the following one, also in two dimensions: $L_{ij} L_{jk} L_{ki} = \frac{3}{2} L_{ii} L_{jk} L_{kj} - \frac{1}{2} L_{ii} L_{jj} L_{kk}$ (we will return to this *reducibility* property below).

At this point an interesting question presents itself: can we single out a finite (preferably minimal) number of invariants that completely characterise the N-jet and in terms of which we can express all other invariants? The answer is affirmative; we will give a constructive proof of such a complete set later on.

Less obvious is the existence, proved by Hilbert [21], of a finite set of *polynomial*, so-called *irreducible invariants*, which can be combined in a polynomial way to express all other polynomial invariants (see also [6], [7], [9]. Because of the restriction to polynomials only, one has to take into account the existence of so-called *syzygies*, i.e., polynomial invariants that need to be supplemented to a minimal set but that are not independent of this. A detailed treatment of this is beyond the scope of this paper. However, as an illustration of Hilbert's theorem, we will give a set of irreducible polynomial invariants for the 2-jet, which holds in arbitrary dimensions. But first we will introduce a *diagrammatic representation* of our polynomial invariants, in the spirit of the Feynman diagrams

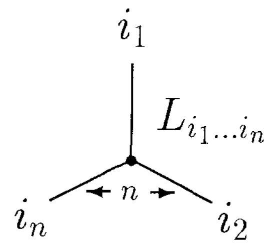


Fig. 1. Diagrammatic representation of the local jet tensor $L_{i_1 \cdots i_n}$ as an n-vertex.

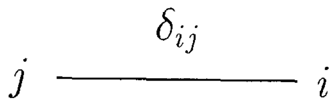


Fig. 2. Diagrammatic representation of the Kronecker tensor δ_{ij}.

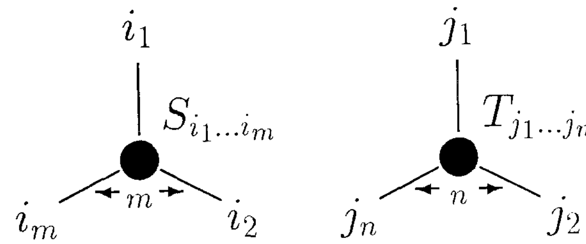


Fig. 3. Tensor product $S_{i_1 \cdots i_m} T_{j_1 \cdots j_n}$. The black disks represent the objects S and T and may consist of arbitrary combinations of basic diagrams with m and n external branches, respectively.

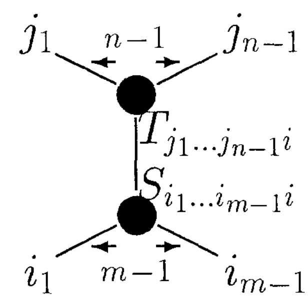


Fig. 4. Tensor contraction $T_{j_1 \cdots j_{n-1} i} S_{i_1 \cdots i_{m-1} i}$

used in particle physics. This is a symbolic representation that allows one to quickly disentangle the various contractions and free indices in an N-jet tensor; it is especially beneficial for more complicated formulas.

Let us now turn to the definition of a diagram. Consider the symmetric n-tensor $L_{i_1 \cdots i_n}$. Diagrammatically, this tensor is represented by an *n-vertex*, i.e., a dot (representing the image L at a given base point and inner scale) with n external branches attached to it, one for each free index; see figure 1. The Kronecker tensor δ_{ij} is represented by a line segment, the ends of which correspond to its two indices; see figure 2. The index symmetry of the tensors $L_{i_1 \cdots i_n}$ and δ_{ij} is manifest in these diagrams. The ε-diagram, finally, must reflect the antisymmetric nature of the Lévi–Civita tensor, and its number of branches must equal the dimension of space. In the two dimensional 2-D case you could use a diagram similar to that for the Kronecker tensor (figure 2) with an extra internal arrow to reflect its orientation, say, pointing from index i to index j in ε_{ij}. By convention, a reversal of internal flow then induces a relative minus sign. We will henceforth restrict ourselves to diagrams corresponding to absolute tensors only.

Having defined these elementary pieces, we can formulate a "diagrammar" on the basis of the theorems of Cartesian tensor calculus. For example, we can form arbitrary tensor products by simply concatenating the individual diagrams, as in figure 3, or we can contract two indices of a tensor by joining the endpoints of two external lines in the diagram, as in figure 4. In this way we can systematically form homogeneous, polynomial N-jet invariants of degree H by forming closed combinations (i.e., without external lines) of H image vertices, each of which has N external lines at the most. Some simple examples taken from the 2-jet are given in figure 5, whereas figure 6 shows some higher-order diagrams. Note that $\delta_{ii} = D$ is a constant (image-independent) invariant, diagrammatically denoted by a single, closed loop.

The reader may verify the following relationship among the numbers of n-vertices, internal and external lines, and closed loops.

Proposition 1 (constraints on tensorial diagrams). Let there be given a homogeneous polynomial (not necessarily connected) diagram with $H = \sum_n V_n$ vertices, where V_n is the num-

Fig. 5. Some basic polynomial 2-jet invariants: L, L_iL_i, $L_iL_{ij}L_j$, L_{ii}, and $L_{ij}L_{ji}$, respectively.

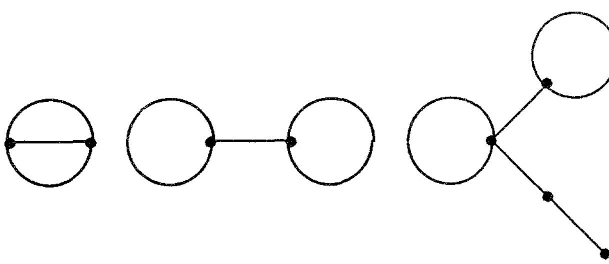


Fig. 6. Some higher-order invariants: $L_{ijk}L_{ijk}$, $L_{iij}L_{jkk}$, and $L_{iijk}L_{jll}L_{km}L_m$. Note that for these more complicated invariants it becomes a cumbersome job to keep track of terms (let alone invariance!) if the Einstein convention is not used, even in the 2-D case: $L_{iijk}L_{jll}L_{km}L_m = (((((L_{xxxx} + L_{xxyy})(L_{xxx} + L_{xyy})) + ((L_{xxxy} + L_{xyyy})(L_{xxy} + L_{yyy})))L_{xx}) + ((((L_{xxxy} + L_{xyyy})(L_{xxx} + L_{xyy})) + ((L_{xxyy} + L_{yyyy})(L_{xxy} + L_{yyy})))L_{xy}))L_x + (((((L_{xxxx} + L_{xxyy})(L_{xxx} + L_{xyy})) + ((L_{xxxy} + L_{xyyy})(L_{xxy} + L_{yyy})))L_{xy}) + ((((L_{xxxy} + L_{xyyy})(L_{xxx} + L_{xyy})) + ((L_{xxyy} + L_{yyyy})(L_{xxy} + L_{yyy})))L_{yy}))L_y$. But even the number of indices in the condensed Einstein summation may be hard to keep track of in more cumbersome expressions. The diagrammatic representation is particularly convenient here: manifest invariance is a plain consequence of the fact that we have a closed diagram.

ber of n-vertices, with E external and I internal lines and with C closed loops. Then for a tensorial diagram of rank E the following constraints hold:

$$\begin{cases} C &= 1 + I - H, \\ \sum_n nV_n &= E + 2I. \end{cases}$$

In particular, we have $E = 0$ for an invariant diagram.

The proof of the existence of a finite set of irreducible polynomial invariants, in terms of which all polynomial invariants can be expressed through multiplication and addition, is a nontrivial result due to Hilbert. Irreducibility means that none of the invariants in the set can be written as a polynomial of the other ones, but it does *not* imply that they are independent, since there generally exist syzygies. A constructive proof is extremely difficult for the general case. If we restrict ourselves to the case of a 2-jet, however, we can state the following proposition.

PROPOSITION 2 (complete set of irreducible polynomial 2-jet invariants). A complete and irreducible system of (absolute) polynomial 2-jet invariants in D dimensions without syzygies is given by all connected diagrams built out of 0-, 1- and 2-vertices with at most D internal lines and no external lines.

Note that figure 5 is just this irreducible set for the 2-D case. One recognises the familiar Laplacian L_{ii} and Canny's "edge detector" L_iL_i [39], but these are put into a context of a *complete* local 2-jet description; together with the other three they unambiguously determine the image's local structure up to second order (inclusive), modulo an arbitrary rotation. Figure 7 shows an example of reducibility. A proof of Proposition 2 can be found in Appendix B.

Finding irreducible sets of polynomial invariants is generally a difficult problem. However, there is a much simpler way to construct complete sets of invariants, albeit nonpolynomial, that at the same time captures all pseudo invariants, viz, by using a system of *gauge coordinates*. We will demonstrate the idea by some explicit examples.

Consider the 2-D case for the sake of simplicity. For a given base point we may introduce local Cartesian coordinates v and w, say, in such a way that $L_v = 0$. This *gauge choice* is always realisable through a suitable Cartesian transformation of an arbitrary Cartesian frame, and hence it is *admissible*. It establishes a frame in which the w-axis is tangent to the image gradient at the origin, and in which the v-axis is tangent to the isophote (i.e., a level curve in two dimensions and a level surface in three dimensions) at the origin.[3] We will agree on taking $L_w = \sqrt{L_iL_i}$ and using a positively oriented, orthonormal basis. Of course, the (v, w)-gauge is ill defined in points with vanishing gradient, but these points form a countable set, at least in a generic image.

Now we can use the following (pseudo) invariant differential operators (the notation is

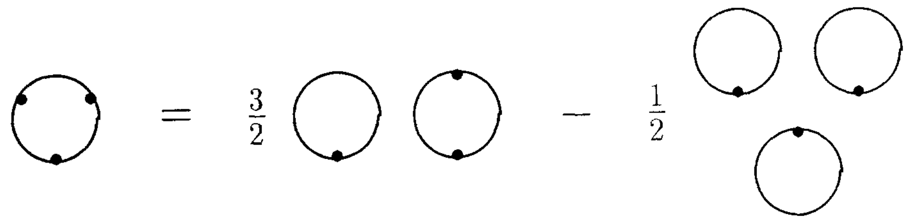


Fig 7. An example of reducibility that holds in two dimensions: $L_{ij}L_{jk}L_{ki} = \frac{3}{2}L_{ii}L_{jk}L_{kj} - \frac{1}{2}L_{ii}L_{jj}L_{kk}$.

self-explanatory).

DEFINITION 3 ((pseudo) invariant differential operators in two dimensions).

$$\partial_v = \frac{1}{\sqrt{L_k L_k}}\varepsilon_{ij}L_j\partial_i,$$
$$\partial_w = \frac{1}{\sqrt{L_k L_k}}\delta_{ij}L_j\partial_i.$$

It is understood that all derivatives L_i in this expression are to be evaluated at the base point of interest, and hence they are considered constant.

Note that v is a pseudo invariant and w an absolute invariant. Consequently, applying the first operator of Definition 3 an odd number of times will yield a pseudo invariant.

Applying these directional differential operators to an image will yield manifest invariants. It is then straightforward to write down a complete set of (pseudo) invariants.

PROPOSITION 3 (complete set of (pseudo) invariants in two dimensions). A complete set of (pseudo) invariants in two dimensions is given in (v, w)-gauge coordinates by the set

$$\mathcal{G} = \left\{ \frac{\partial^{m+n}L}{\partial v^m \partial w^n} \right\}_{(m,n)\neq(1,0)},$$

in which the indices (m, n) run over all possible orders up to a desired jet order.

For many lowest-order geometrical issues the vanishing of L_v may greatly facilitate theoretical manipulations without any loss of generality (for we can always use Definition 3 to recast expressions in (v, w)-gauge back to arbitrary coordinates). To illustrate this, we consider the Laplacian L_{ii}. We may decompose this Laplacian as $L_{ii} = L_{vv} + L_{ww}$, showing that its zero crossings deviate from gradient extrema (edges and phantom edges corresponding to $L_{ww} = 0$) by an amount determined by the invariant L_{vv}. To give this a geometrical interpretation, consider the image's isophote passing through the base point of interest, $(v, w) = (0, 0)$, i.e., the curve defined implicitly by $L = L_0$. Taking the (total) first- and second-order derivatives of this with respect to v yields $L_w w' + L_v = 0$ (a prime denotes a derivative d/dv) and $L_{vv} + 2L_{vw}w' + L_{ww}w'^2 + L_w w'' = 0$, from which we may solve for the intrinsic isophote properties $w' = 0$ and $w'' = -L_{vv}/L_w$ in the point of reference. But w'' is just the *isophote curvature* κ, so that we have obtained the result that the Laplacian zero-crossings may correspond to edges (or phantom edges) only if the isophotes are locally sufficiently straight (relative to the inner scale): $L_{ii} = L_{ww} - \kappa L_w$. This well-known result merely serves as an illustration of the use of gauge coordinates tuned to a particular geometrical problem (see also [40]–[42]).

Of course, other gauges are possible. If somehow second-order structure is emphasised, it may be beneficial to impose a gauge condition in which the mixed second-order derivative vanishes, $L_{pq} = 0$, say. This gauge is admissible since it is always possible to diagonalise the symmetric Hessian L_{ij} by a suitable rotation. We can then replace the set $\mathcal{G}$ in Proposition 3 by the following set.

PROPOSITION 4 (complete set of (pseudo) invariants in two dimensions). A complete set of (pseudo) invariants in two dimensions is given in (p, q)-gauge coordinates by the set

$$\mathcal{H} = \left\{ \frac{\partial^{m+n} L}{\partial p^m \partial q^n} \right\}_{(m,n) \neq (1,1)},$$

in which the indices (m, n) run over all possible orders up to a desired jet order.

Similar gauges apply to higher-dimensional cases, e.g., in three dimensions we may introduce gauge coordinates (u, v, w) by the gauge conditions $L_u = L_v = L_{uv} = 0$ ($L_w = \sqrt{L_i L_i}$, (u, v, w) positively oriented) or gauge coordinates (p, q, r) associated with a pure second-order gauge $L_{pq} = L_{pr} = L_{qr} = 0$. More generally, in a D-dimensional image we may introduce gauge coordinates $(u_1, \ldots, u_{D-1}, w)$ by the $D - 1$ gauge conditions $L_{u_a} = 0 (a = 1, \ldots, D - 1)$ and the $\frac{1}{2}(D-1)(D-2)$ gauge conditions $L_{u_a u_b} = 0 (a \neq b = 1, \ldots, D - 1)$, making up a total of $\frac{1}{2}D(D-1)$ equations, i.e., the same as for a system of gauge coordinates $(p_1, \ldots, p_D)$ associated with a pure second-order gauge $L_{p_i p_j} = 0 (i \neq j = 1, \ldots, D)$.

To illustrate the theoretical convenience of gauge coordinates once more, let us derive an expression for the isophote curvatures in the three-dimensional (3-D) case. In a 3-D image a generic isophote is a (closed) surface. The lowest-order deviation from its tangent plane at a given point can be expressed by two independent invariants, viz., the *principal curvatures* or, alternatively, the *mean* and the *Gaussian curvature*. Recall the 2-D case, for which the (single) isophote curvature κ was given by $\kappa = -L_{vv}/L_w$ in the analogous (v, w)-gauge. For the 3-D case we can obtain a curve by taking a normal section of the isophote surface. Using the rotational degree of freedom around the surface normal (i.e., the gradient direction), we can measure its curvature as a function of the rotation angle, yielding the so-called *Dupin's indicatrix*. This is a cone representing the surface's normal curvature (or rather, the reciprocal of the square root of its absolute value) as a function of angle. The two axes of the indicatrix correspond to the principal directions, in which the normal curvatures reach extremal values: the principal curvatures. One can show that the u and v axes are tangent to the principal directions. If the surface is locally umbilical (i.e., if the two principal curvatures are equal) or if the gradient vanishes, the (u, v, w)-gauge is ill defined, but again for a generic image, this happens only at isolated points. Now it is clear from the 2-D case that in the (u, v, w)-gauge the principal curvatures are given by $\kappa_1 = -L_{uu}/L_w$ and $\kappa_2 = -L_{vv}/L_w$, whereas the mean curvature h and Gaussian curvature k are given by their average and product, respectively: $h = \frac{1}{2}(\kappa_1 + \kappa_2)$, $k = \kappa_1 \kappa_2$. To arrive at the expression in a general Cartesian frame, simply write down a rational manifest invariant with the appropriate degrees of homogeneity for its first- and second-order derivatives. One can make a guess—with a modest amount of foresight—among the few simplest second-order invariants one can imagine that meet the appropriate homogeneity constraints

$$h = \frac{1}{2} \frac{L_i L_{ij} L_j - L_i L_i L_{jj}}{(L_k L_k)^{3/2}}, \tag{2}$$

$$k = \frac{1}{2} \frac{\varepsilon_{ijk} \varepsilon_{lmn} L_i L_l L_{jm} L_{kn}}{(L_p L_p)^2}. \tag{3}$$

Indeed, evaluating the expressions in gauge coordinates significantly reduces the number of effective terms and precisely yields the expressions that hold in the (u, v, w)-gauge. By virtue of invariance, (2) and (3) apparently do represent the mean and Gaussian curvatures in an arbitrary frame. Note that we did not even need to calculate the principal directions as such (this requires a more involved calculation; see [43]). The importance of isophote properties is underlined by their invariance under a more general transformation group, viz., the product group of the Cartesian (spatial) group and the *group of general intensity transformations* [44], the latter one of which consists of all one-to-one image grey-value maps $L \mapsto \Lambda(L)$ (gamma corrections, contrast and brightness adjustments, etc.).

The reason for casting simple expressions, derived in a specific gauge, back to more complex ones that hold in arbitrary coordinate systems is a practical one: to be able to perform computer calculations on rastered images with respect to any user-defined Cartesian frame. We have already encountered many examples of differential invariants that illustrate our main theory. In section 4 we present the results of applying some of these, as well as some new ones, on test images.

4 More Examples

On the basis of our theory on differential invariants we have built a software implementation for calculating invariants. The main program is designed to parse an expression entered either in index notation, as in "`Lij * Lij`", or in gauge coordinates, as in "`Lvv * Lw^2`", and then to calculate these for a specified range of scales and a given input image. The result of the parsing stage is an executable invariant, which can be called with the given image as its operand. It manipulates (points to) the derivative images required for the invariant. These derivatives reside either in memory or on disk. They are obtained in Fourier space in the straightforward way, viz., by Fourier inversion of the product of the Fourier-transformed image and the appropriate Gaussian kernel, the latter of which is calculated directly in Fourier space (see Definition 2). The advantage of this is the fact that differentiation becomes diagonal in Fourier space.

The boundary problem may effectively be dealt with by a proper downsampling scheme, which establishes a grid for a σ-scaled image with a voxel volume proportional to σ^D. The grid points may be taken to be the centers of a collection of compact σ-neighbourhoods with a fixed relative overlap. We leave open the question of how to choose the proportionality constant k that defines the effective radius $R = k\sigma$ of these compact neighbourhoods (to constrain this *scale-independent* unknown is what the boundary problem actually boils down to). If such a grid is imposed, the boundary grid points are located a distance R away from the image outline.

For the sake of presentation, we have not downsampled the images according to their inner scales. Accordingly, one must ignore the information in a boundary strip proportional to the inner scale of each image. With our Fourier-space technique this information is unreliable because of the well-known wrap-around effect, but it should be appreciated that it is unreliable with any other technique as well; there is no hope for extracting reliable spatial information at inner scale σ in such a boundary strip (without the use of *a priori* knowledge).

The 2-D test images are shown in figure 8, and some examples of first, second, third, and fourth order generated in this way are shown in figures 9–13, respectively. The input images for figures 9–12 were created artificially as binary-valued grey-level images (see figure 8) and were distorted by a fair amount of additive, pixel-uncorrelated Gaussian noise before the evaluation of the invariants. Figure 13 shows an example of an invariant for a typical 2-D slice (for the sake of presentation) taken from a 3-D medical data set, i.e. an MRI image.

The figure captions contain all case-specific details.

5. Conclusion and Discussion

We have proposed a theory for the systematic study of local image structure based on tensor calculus and differential invariants. We have defined a simple diagrammar for a pictorial representation of tensors and invariants. This allows one to immediately grasp the manifest invariant structure of such objects without the need for dummy indices or an explicit choice of gauge. A significant part of the well-established mathematical results on differential—algebraic or geometric—invariants has been made operational by the introduction of well-posed scale-space differential operators. High-order differential structure of an image can be revealed despite noise or discretisation artefacts, the criteria of practical interest being the ratio of pixel scale or noise correlation width to the inner scale of differentiation, the ratio of inner scale to the scale of the image's field of view, and the relative grey-level resolution or noise characteristics. All of these factors contribute to the quality of image derivatives. A quantification of this notion of quality that takes into account all of these trade-offs is still missing, although a partial solution has been given by Blom [41], who studied the behaviour of the scale-space differential operators in the presence of additive (both

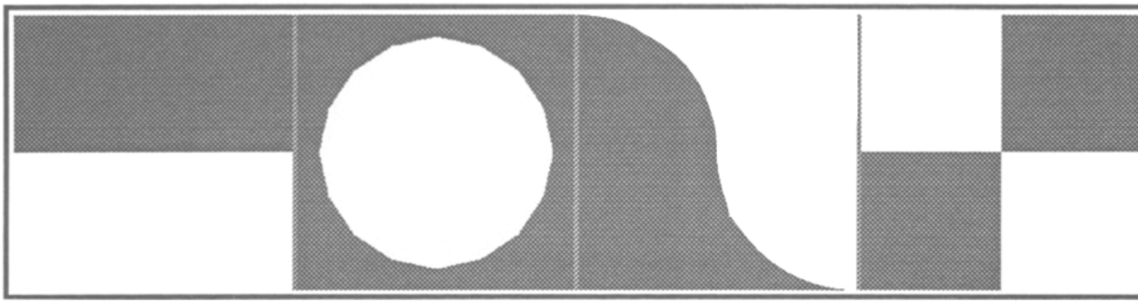


Fig. 8. The 2-D test images used in the examples, before noise perturbation. From left to right: a straight step edge, a filled polygon with 16 corners, an inflected step edge, and a checkerboard pattern. Image dimensions: 256 by 256 pixels.

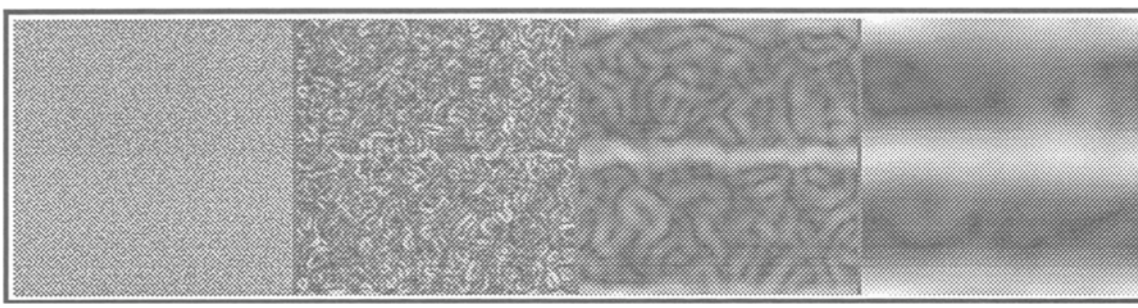


Fig. 9. A first-order differential invariant calculated for an artificially created 2-D test image showing a binary step edge of 100 arbitrary units, perturbed by additive, pixel-uncorrelated Gaussian noise with a standard deviation of 300 units. The left image is the input image used. The other images show the invariant $L_w = \sqrt{L_i L_i}$ for a triple of scales $\ln \sigma = 1, 2, 3$,. As can be seen, with this kind of noise the edge is well represented only at sufficiently l- ge scales.

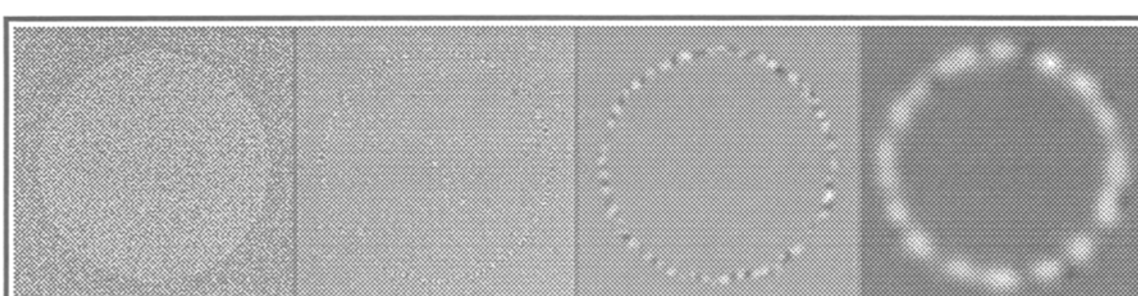


Fig. 10. The invariant $-L_{vv}L_w^2 = L_i L_{ij} L_j - L_i L_i L_{jj}$ calculated for a polygon with 16 corners. The input image shown on the left was created as a binary test image with an intensity difference of 100 units, which was then perturbed by adding pixel-uncorrelated Gaussian noise with a standard deviation of 100 units. This invariant expresses a shear-invariant trade-off between edge-strength (the factor L_w^3) and isophote curvature (the factor $-L_{vv}/L_w$) and therefore could be called "corner-strength." For an in-depth discussion of this invariant the reader is referred to [41]. Scales: $\ln \sigma = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}$. A further increase of scale (not shown here) would make the 16 corners merge together along the disk boundary, creating a circumference of more or less uniform corner strength, corresponding to the uniform isophote curvature and edge strength at the boundary of the approximating disk that survives at those levels of scales. This phenomenon, by the way, i.e., the small-scale existence of the individual corners in this example, qualitatively accounts for the fact that a further increase of noise would make the corners inaccessible *at any scale*.

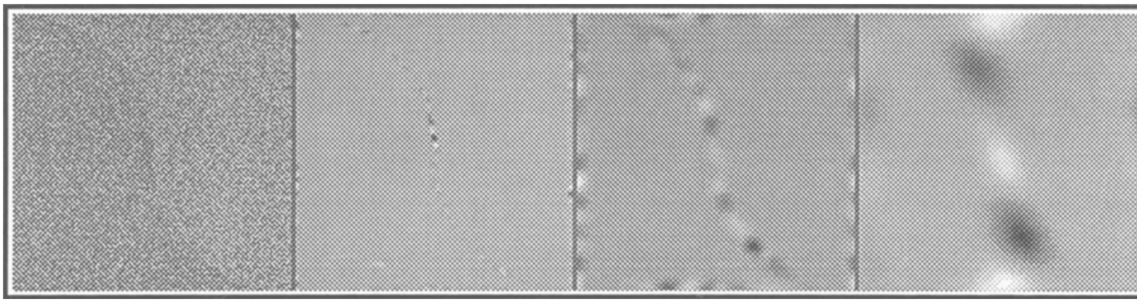


Fig. 11. The invariant $L_{vvv}L_w^5 - 3L_{vv}L_{vw}L_w^4$, or $3(L_{yy}L_x^2 - 2L_yL_xL_{xy} + L_y^2L_{xx})(L_yL_{yy}L_x - L_y^2L_{xy} + L_x^2L_{xy} - L_yL_xL_{xx}) - (L_y^2 + L_x^2)(L_{yyy}L_x^3 - 3L_yL_x^2L_{xyy} + 3L_y^2L_xL_{xxy} - L_y^3L_{xxx})$, calculated for a noise-perturbed, inflected step edge. The input image shown on the left was created as a binary test image with an intensity difference of 100 units, which was then perturbed by adding pixel uncorrelated Gaussian noise with a standard deviation of 150 units. This third-order invariant measures a trade-off between edge strength (the factor L_w^6) and rate of change of isophote curvature along the isophote (the remainder) and therefore could be called "inflexion strength." Scales: $\ln \sigma = \frac{3}{2}, \frac{5}{2}, \frac{7}{2}$.

pixel-correlated and pixel-uncorrelated) Gaussian noise as a function of scale.

Many of the classical operators used in computer vision, such as Canny's "edge detector" and the familiar Laplacian zero-crossings, are nicely embedded into the theory, being based on the few simplest differential invariants of lowest orders. Note that these kinds of operators *define* structures (such as edges and zero crossings) i.e., these structures can be identified only by virtue of their defining operators. A complete description of local image structure (up to some order), e.g., given in terms of a complete set of invariants, is the preferred way—not ad hoc by virtue of completeness—to extract input data for further processing in high-level image routines.

Appendix A: Cartesian Tensors

In this appendix we review some basic definitions and results from Cartesian tensor calculus (for an easy introduction, also [47], [48]). A Cartesian transformation in a (real) Cartesian D-dimensional space V^D is defined by a translation and an orthogonal transformation of the Cartesian coordinates of V^D:

$$\widetilde{x}_i = a_{ij}x_j + b_i, \tag{4}$$

with a_{ij} satisfying the orthogonality constraint.

$$a_{ki}a_{kj} = a_{ik}a_{jk} = \delta_{ij}. \tag{5}$$

Here, and in all that follows, all Latin indices run from 1 to D and the Einstein summation convention is in effect, i.e., repeated indices in a tensor product imply a summation over all possible values: $T_{ii} \stackrel{\text{def}}{=} \sum_{i=1}^{D} T_{ii}$. Furthermore, δ_{ij} is the Kronecker symbol defined as usual: $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise. Inversion of (4), by using (5), yields

$$x_i = a_{ji}(\widetilde{x}_j - b_j), \tag{6}$$

and so we have

$$\frac{\partial \widetilde{x}_i}{\partial x_j} = a_{ij}, \quad \frac{\partial x_i}{\partial \widetilde{x}_j} = a_{ji}. \tag{7}$$

A Cartesian p-tensor can be defined with respect to some arbitrary Cartesian coordinate system as a D^p-tuple of real numbers $C_{i_1 \cdots i_p}$ with a particular transformation property, given in the following definition.

DEFINITION 4. A D^p-tuple of real numbers $C_{i_1 \cdots i_p}$ is called a (Cartesian) tensor of rank p (or p-tensor) if a Cartesian coordinate transformation $\widetilde{x}_i = a_{ij}x_j + b_i$ induces the following transformation of its components:

$$\widetilde{C}_{i_1 \cdots i_p} = a_{i_1 j_1} \cdots a_{i_p j_p} C_{j_1 \cdots j_p}. \tag{8}$$

In particular, we distinguish a 0-tensor, or *scalar*, and a 1-tensor, or *vector* (since we are only interested in Cartesian tensors, we

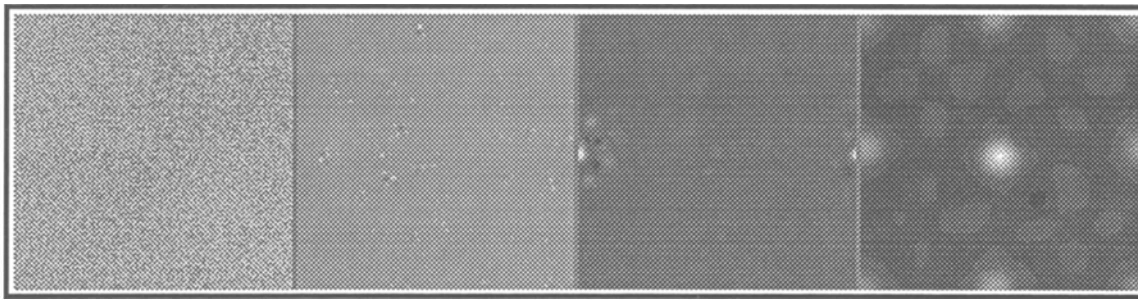


Fig. 12. The pure fourth-order invariant $D = -I^3 + 27J^2$ with $I = \frac{1}{2}L_{iijj}L_{kkll} - L_{iijk}L_{jkll} + \frac{1}{2}L_{ijkl}L_{ijkl}$ and $J = \frac{1}{8}L_{iijj}L_{kkll}L_{mmnn} - \frac{1}{8}L_{iijj}L_{klmn}L_{klmn} - \frac{1}{4}L_{iijj}L_{kklm}L_{lmnn} + \frac{1}{4}L_{iijk}L_{jklm}L_{lmnn}$, or $D = -1(L_{xxxx}L_{yyyy} - 4L_{xxxy}L_{xyyy} + 3L_{xxyy}L_{xxyy})^3 + 27(L_{xxxx}(L_{xxyy}L_{yyyy} - L_{xyyy}L_{xyyy}) - L_{xxxy}(L_{xxxy}L_{yyyy} - L_{xxyy}L_{xyyy}) + L_{xxyy}(L_{xxxy}L_{xyyy} - L_{xxyy}L_{xxyy}))^2$, calculated for a checkerboard pattern. The input image shown on the left was created as a binary test image with an intensity difference of 100 units, which was then perturbed by adding pixel-uncorrelated Gaussian noise with a standard deviation of 300 units. Scales: $\ln\sigma = \frac{3}{2}, \frac{5}{2}, \frac{7}{2}$. This invariant expresses an algebraic property of a fourth-order binary form $f(\mathbf{x}) = \frac{1}{24}L_{ijkl}x_i x_j x_k x_l$ and is called the (fourth-order) *discriminant*. It is a well-known invariant in the mathematical literature on algebraic invariants, usually expressed in a bracket formalism in terms of so-called *transvectants*. For a more detailed description of this the reader is referred to [15]. Note that despite the high amount of noise this fourth-order property is well represented at an appropriate scale.

will henceforth omit the adjective "Cartesian"). If the $C_{i_1 \cdots i_p}$ are functions of the coordinates $\mathbf{x} = (x_1, \ldots, x_D)$, then we call $C_{i_1 \cdots i_p}(\mathbf{x})$ a *tensor field* (and, in particular, a *scalar field* if $p = 0$ and a *vector field* if $p = 1$).

A tensor A_{ij} is called *symmetric* with respect to i and j if $A_{ij} = A_{ji}$, and it is called *antisymmetric* to i and j if $A_{ij} = -A_{ji}$. The generalisation to arbitrary p-tensors is obvious.

Readers may convince themselves that the Kronecker symbol defined above defines an invariant tensor with constant components, justifying the following suggestive index notation.

CLAIM 1 (Kronecker tensor). The Kronecker symbol δ_{ij} is an invariant, symmetric tensor.

The Kronecker tensor is also often referred to as the *fundamental tensor.*

DEFINITION 5 (contraction). A contraction is an operation that reduces a p-tensor ($p \geq 2$) to a ($p-2$)-tensor by equating (and hence, by convention, summing over) a pair of free indices. So if $A_{i_1 \cdots i_p}$ is a p-tensor, then a contraction on i_j and $i_k (1 \leq j < k \leq p)$ yields $A_{i_1 \cdots i_{j-1} i i_{j+1} \cdots i_{k-1} i i_{k+1} \cdots i_p}$.

By transforming a contracted tensor according to (8) and using the orthogonality constraint (5) it is shown that a contracted tensor is indeed a tensor.

An important observation is the following.

CLAIM 2. The nth order partial differential operator $\partial^n / \partial x_{i_1} \cdots \partial x_{i_n}$ formally transforms as an n-tensor.

This property is unique to the Cartesian group and is easily proved by performing a Cartesian coordinate transformation in which the chain rule and the orthogonality property are used (5).

Of course, the nth-order partial differential operator in Claim 2 is the formal product of n first-order (gradient) operators. In general we have the following claim.

CLAIM 3. If $A_{i_1 \cdots i_p}$ and $B_{i_1 \cdots i_q}$ are Cartesian p- *and* q-tensors, respectively, then the *tensor product*, defined by

$$P_{i_1 \cdots i_{p+q}} = A_{i_1 \cdots i_p} B_{i_{p+1} \cdots i_{p+q}},$$

is a $(p + q)$-tensor.

The proof is again based on straightforward transformation. In a similar way we may prove the following tensor properties.

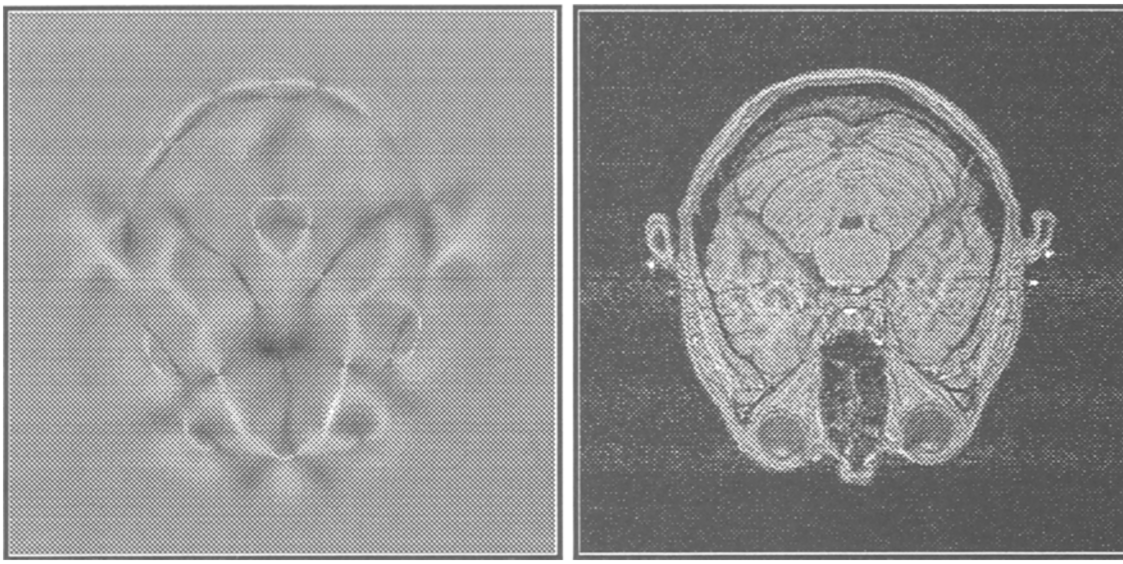


Fig. 13. The second-order invariant $-L_{vv}$ (right image) calculated for the 256 x 256 MRI image shown on the left at scale $\ln \sigma = 2$ (with σ given in pixel units). This invariant is a measure of the rate of change of the image gradient direction along the isophote (to see this, just note that isophote curvature equals $\kappa = -L_{vv}/L_w$, as explained in the text, so that $-L_{vv} = \kappa L_w$ is the isophote curvature weighted by the gradient magnitude). For this reason it can be used to find ridgelike structures in the image. This invariant and a 3-D generalisation of it have been applied successfully for matching CT and MRI data; see [45], [46].

CLAIM 4. If $A_{i_1 \cdots i_p}$ and $B_{i_1 \cdots i_p}$ are Cartesian p-tensors, then so is any linear combination $C_{i_1 \cdots i_p} = \lambda A_{i_1 \cdots i_p} + \mu B_{i_1 \cdots i_p}$ in which λ and μ are scalars.

CLAIM 5. Symmetry (antisymmetry) is preserved under Cartesian coordinate transformations, i.e., if $A_{ij} \stackrel{\text{def}}{=} A_{i_1 \cdots i_{j-1} i i_{j+1} \cdots i_{k-1} j i_{k+1} \cdots i_p}$, then $A_{ij} = (-)A_{ji}$ implies that also $\widetilde{A}_{ij} = (-)\widetilde{A}_{ji}$.

CLAIM 6. The validity of a tensor equation in all Cartesian coordinate systems follows from the following lemma: if $A_{i_1 \cdots i_p} = 0$, then also $\widetilde{A}_{i_1 \cdots i_p} = 0$.

Put differently: if $A_{i_1 \cdots i_p} = B_{i_1 \cdots i_p}$, then also $\widetilde{A}_{i_1 \cdots i_p} = \widetilde{B}_{i_1 \cdots i_p}$.

A useful generalisation of tensor calculus also comprises so-called *pseudo tensors* (sometimes called *tensor densities* or *relative tensors*).

DEFINITION 6. A D^p-tuple of real numbers $U_{i_1 \cdots i_p}$ is called a pseudo tensor of rank p if a Cartesian coordinate transformation $\widetilde{x}_i = a_{ij} x_j + b_i$ induces the following transformation of its components:

$$\widetilde{U}_{i_1 \cdots i_p} = \Delta a_{i_1 j_1} \cdots a_{i_p j_p} U_{j_1 \cdots j_p}, \qquad (9)$$

in which $\Delta = \pm 1$ is the determinant of the orthogonal matrix a_{ij}.

An important *invariant pseudo tensor* is given by the *Lévi–Civita tensor* (the attribute "pseudo" is often omitted). It is defined by normalising one of its components to unity, $\varepsilon_{1 \ldots D} \stackrel{\text{def}}{=} 1$, and fixing the other $D^D - 1$ components by requiring it to be completely antisymmetric with respect to any pair of indices. In other words, $\varepsilon_{i_1 \cdots i_D}$ equals the sign of the permutation of the indices $(i_1 \cdots i_D)$. The following claim shows that this is indeed a good definition.

CLAIM 7 (Lévi-Civita pseudo tensor). The completely antisymmetric Lévi–Civita symbol $\varepsilon_{i_1 \cdots i_D}$, defined uniquely by the normalisation $\varepsilon_{1 \ldots D} \stackrel{\text{def}}{=} 1$, is an invariant pseudo tensor.

The proof of this is perhaps somewhat less obvious than those of all previous claims, and so we will present it here: transform the single independent component $\tilde{\varepsilon}_{1\ldots D} = \Delta a_{1i_1} \cdots a_{Di_D} \varepsilon_{i_1 \cdots i_D} = \Delta^2 = 1$ (by definition of the determinant of a_{ij}). Since antisymmetry is preserved, this means that we have $\tilde{\varepsilon}_{i_1 \cdots i_D} = \varepsilon_{i_1 \cdots i_D}$, which completes the proof. Note that the Lévi–Civita symbol is well defined as a true tensor if we consider the Cartesian subgroup connected to the identity, for which we always have $\Delta = 1$. Apparently a pseudo tensor describes an object that has an intrinsic orientation (like a left or right hand).

The following are some useful and easily verifiable results relating to pseudo tensors.

CLAIM 8. Let A denote a true tensor, and let U and V denote pseudo tensors. Then the following results hold:

a) Any linear combination $W_{i_1 \cdots i_p} = \lambda U_{i_1 \cdots i_p} + \mu V_{i_1 \cdots i_p}$, in which λ and μ are scalars, is a pseudo tensor.
b) The tensor product $W_{i_1 \cdots i_{p+q}} = A_{i_1 \cdots i_p} U_{i_{p+1} \cdots i_{p+q}}$ is a pseudo tensor.
c) The tensor product $B_{i_1 \cdots i_{p+q}} = U_{i_1 \cdots i_p} V_{i_{p+1} \cdots i_{p+q}}$ is a tensor.
d) A contracted pseudo tensor is a pseudo tensor.

Particularly useful pseudo tensors are obtained by forming products of a true tensor and the Lévi–Civita pseudo tensor. A full contraction of a D-tensor onto the ε-tensor is also called an *alternation*. A full contraction of all indices in a tensor yields an invariant (which is pseudo if and only if the number of ε-tensors involved is odd).

According to Claim 8 the product of two pseudo tensors is a true tensor, and so the product of two constant ε-tensors apparently yields a constant true tensor. It can be shown that any constant true tensor can be written as a linear combination of products of the fundamental δ-tensor. Indeed, we have the following result.

DEFINITION 7. The generalised Kronecker tensor of rank $2n (n = 0, \cdots D)$ is defined by

$$\delta_{\mu_1 \cdots \mu_n ; \nu_1 \cdots \nu_n} = \det A_{\mu_1 \cdots \mu_n ; \nu_1 \cdots \nu_n},$$

in which the $n \times n$ matrix $A_{\mu_1 \cdots \mu_n ; \nu_1 \cdots \nu_n}$ is given by

$$A_{\mu_1 \cdots \mu_n ; \nu_1 \cdots \nu_n} = \begin{pmatrix} \delta_{\mu_1 \nu_1} & \cdots & \delta_{\mu_1 \nu_n} \\ \vdots & & \vdots \\ \delta_{\mu_n \nu_1} & \cdots & \delta_{\mu_n \nu_n} \end{pmatrix}.$$

This determinant is an $n!$-sum of n-products of the fundamental Kronecker tensor, and so it indeed defines a tensor of rank $2n$. By definition, this tensor is antisymmetric with respect to its first n indices, as well as with respect to its last n indices. For $n = 0$ we define $\delta = A = 1$, and for $n = 1$ we indeed regain the familiar Kronecker tensor.

CLAIM 9. A double product of ε-tensors can be written as a polynomial of δ-tensors:

$$\varepsilon_{\mu_1 \cdots \mu_n \lambda_1 \cdots \lambda_{D-n}} \varepsilon_{\nu_1 \cdots \nu_n \lambda_1 \cdots \lambda_{D-n}} = (D-n)! \delta_{\mu_1 \cdots \mu_n ; \nu_1 \cdots \nu_n}.$$

Proof. We may embed the $n \times n$ matrix $A_{\mu_1 \cdots \mu_n ; \nu_1 \cdots \nu_n}$, introduced in Definition 7, into a $D \times D$-matrix $\tilde{A}^{[D]}_{\mu_1 \cdots \mu_n ; \nu_1 \cdots \nu_n}$, which has the same determinant, by adding a $(D-n) \times (D-n)$ identity block $\mathbf{1}$, as follows.

DEFINITION 8.

$$\tilde{A}^{[D]}_{\mu_1 \cdots \mu_n ; \nu_1 \cdots \nu_n} = \left(\begin{array}{c|c} \tilde{A}_{\mu_1 \cdots \mu_n ; \nu_1 \cdots \nu_n} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{1} \end{array} \right).$$

Using its standard definition, we can write the determinant of the matrix $\tilde{A}^{[D]}_{\mu_1 \cdots \mu_n ; \nu_1 \cdots \nu_n}$ with the use of a double ε-product:

$$\det \tilde{A}^{[D]}_{\mu_1 \cdots \mu_n ; \nu_1 \cdots \nu_n} = \frac{1}{D!} \varepsilon_{\alpha_1 \cdots \alpha_D} \varepsilon_{\beta_1 \cdots \beta_D} \times (\tilde{A}^{[D]}_{\mu_1 \cdots \mu_n ; \nu_1 \cdots \nu_n})_{\alpha_1 \beta_1} \cdots \times (\tilde{A}^{[D]}_{\mu_1 \cdots \mu_n ; \nu_1 \cdots \nu_n})_{\alpha_D \beta_D}.$$

Despite its D^{2D} terms, this is actually a relatively sparse sum: only those terms for which the indices $(\alpha_1, \ldots, \alpha_D)$ and $(\beta_1, \ldots, \beta_D)$ are permutations of $(1, \ldots, D)$ survive, and we may reorder them in such a way that the first n indices address the actual matrix elements of $A_{\mu_1\cdots\mu_n;\nu_1\cdots\nu_n}$. To this end we consider all permutations $(k_1, \ldots, k_D)$ of $(1, \ldots, D)$ such that the n-tuple $(\alpha_{k_1}, \ldots, \alpha_{k_n})$ and the $(D-n)$-tuple $(\alpha_{k_{n+1}}, \ldots, \alpha_{k_D})$ are permutations of $(1, \ldots, n)$ and $(n+1, \ldots, D)$, respectively. If we take into account a combinatorial factor, counting the various possibilities to choose this index separation (all of which give equal contributions), we may assume that the indices $\alpha_{k_1}, \ldots, \alpha_{k_n}$ and $\beta_{k_1}, \ldots, \beta_{k_n}$ index the actual matrix elements of the block $A_{\mu_1,\ldots,\mu_n;\nu_1,\ldots,\nu_n}$, whereas $\alpha_{k_{n+1}}, \ldots, \alpha_{k_D}$ and $\beta_{k_{n+1}}, \ldots, \beta_{k_D}$ index the identity block. In other words, $(\widetilde{A}^{[D]}_{\mu_1\cdots\mu_n;\nu_1\cdots\nu_n})_{\alpha_{k_i}\beta_{k_i}} = (A_{\mu_1\cdots\mu_n;\nu_1\cdots\nu_n})_{\alpha_{k_i}\beta_{k_i}} = \delta_{\mu_{\alpha_{k_i}}\nu_{\beta_{k_i}}}$, for the first n subindices $i = 1, \ldots, n$, and $(\widetilde{A}^{[D]}_{\mu_1\cdots\mu_n;\nu_1\cdots\nu_n})_{\alpha_{k_i}\beta_{k_i}} = \delta_{\alpha_{k_i}\beta_{k_i}}$, for the last $D-n$ indices $i = n+1, \ldots, D$. This leads to the following expression:

$$\det\widetilde{A}^{[D]}_{\mu_1\cdots\mu_n;\nu_1\cdots\nu_n} = \binom{D}{n}\frac{1}{D!}\varepsilon_{\alpha_{k_1}\cdots\alpha_{k_D}} \times\varepsilon_{\beta_{k_1}\cdots\beta_{k_D}}\delta_{\mu_{\alpha_{k_1}}\nu_{\beta_{k_1}}}\cdots \times\delta_{\mu_{\alpha_{k_n}}\nu_{\beta_{k_n}}}\delta_{\alpha_{k_{n+1}}\beta_{k_{n+1}}}\cdots \times\delta_{\alpha_{k_D}\beta_{k_D}},$$

or, on a relabeling of dummy contraction indices,

$$\det\widetilde{A}^{[D]}_{\mu_1\cdots\mu_n;\nu_1\cdots\nu_n} = \frac{1}{n!(D-n)!}\varepsilon_{\alpha_{k_1}\cdots\alpha_{k_n}\lambda_1\cdots\lambda_{D-n}} \times\varepsilon_{\beta_{k_1}\cdots\beta_{k_n}\lambda_1\cdots\lambda_{D-n}}\delta_{\mu_{\alpha_{k_1}}\nu_{\beta_{k_1}}}\cdots \times\delta_{\mu_{\alpha_{k_n}}\nu_{\beta_{k_n}}}.$$

For a given n-tuple $(k_1, \ldots, k_n)$ there are $n!$ equally contributing terms in this expression (corresponding to all permutations of this n-tuple), and so we may finally rewrite this expression as

$$\det\widetilde{A}^{[D]}_{\mu_1\cdots\mu_n;\nu_1\cdots\nu_n} = \frac{1}{(D-n)!}\varepsilon_{\mu_1\cdots\mu_n\lambda_1\cdots\lambda_{D-n}} \times\varepsilon_{\nu_1\cdots\nu_n\lambda_1\cdots\lambda_{D-n}}.$$

Since $\det A_{\mu_1\cdots\mu_n;\nu_1\cdots\nu_n} = \det \widetilde{A}^{[D]}_{\mu_1\cdots\mu_n;\nu_1\cdots\nu_n}$, we have completed the proof of Claim 9.

Appendix B: Proof of Irreducibility

In this appendix we give a proof of the irreducibility of the system of polynomial invariants given in Proposition 2, i.e., the set $\{L, S_0, \ldots, S_{D-1}, I_1, \ldots, I_D\}$, with $S_n = L_{i_1}L_{i_1i_2}L_{i_2i_3}\cdots L_{i_ni_{n+1}}L_{i_{n+1}}$ and $I_n = L_{i_1i_2}L_{i_2i_3}\cdots L_{i_ni_1}$ (both S_n and I_n contain n 2-vertices). Note that all connected polynomial diagrams are of the form L, S_n, or I_n for some $n = 0, 1, 2, \ldots$.

We will first consider a system with 2-vertices only and show that all I_n for $n > D$ are reducible. Then we will turn to the general case, first by including the S_n and showing their reducibility for $n > D-1$, and then by extending it with the trivial zeroth-order member L.

B.1 Irreducible System for $\{L_{ij}\}$

We concentrate on the second-order system $\{L_{ij}\}$. Consider the following definitions (see also Definition 7).

DEFINITION 9.

$$L^{[k]}_{ij} \stackrel{\text{def}}{=} \delta_{i_1\cdots i_k i;j_1\cdots j_k j}L_{i_1j_1}\cdots L_{i_kj_k},$$
$$X^{[k]} \stackrel{\text{def}}{=} \delta_{i_1\cdots i_k;j_1\cdots j_k}L_{i_1j_1}\cdots L_{i_kj_k}.$$

By developing the $(k+1)\times(k+1)$ determinant underlying the generalised Kronecker tensor of rank $k+1$ in the first definition (see Definition 7) with respect to the last column into $k\times k$ determinants, one may derive the following identity:

$$L^{[k]}_{ij} = \delta_{ij}X^{[k]} - kL^{[k-1]}_{i\alpha}L_{\alpha j}. \qquad (10)$$

By induction we then have

$$L^{[k]}_{ij} = (-1)^k k! L^k_{ij} + \sum_{i=1}^{k}\binom{k}{i}X^{[i]}L^{[k-i]}_{ij}, \qquad (11)$$

and since, by construction, the generalised Kronecker tensor of rank $D+1$ vanishes identically, we have found the *Hamilton–Cayley identity*:

$$L^{[D]}_{ij} = 0. \qquad (12)$$

The left-hand side of (12) is a polynomial of order D in the Hessian L_{ij}. This so-called "characteristic polynomial" has D (generally distinct) roots λ_n, $n = 1, \ldots, D$, the eigenvalues of the Hessian, which are functions of the invariants $X^{[k]}$. These D invariants correspond to the D independent degrees of freedom of L_{ij}. Instead of $X^{[n]}$ (or λ_n) we may use the traces I_n, which completes the proof of the irreducibility of the set $\{I_1, \ldots, I_D\}$.

B.2 Irreducible System for $\{L, L_i, L_{ij}\}$

The irreducibility of the set $\{S_0, \ldots, S_{D-1}, I_1, \ldots, I_D\}$ associated with the first- and second-order tensors $\{L_i, L_{ij}\}$ could be proved in a way similar to that the for the second-order case. However, it is more economical to proceed differently, by using previous results.

We introduce two independent parameters λ and μ and consider the following symmetric 2-tensor.

DEFINITION 10.

$$H_{ij}(\lambda, \mu) \stackrel{\text{def}}{=} \lambda L_i L_j + \mu L_{ij}.$$

We can use $H_{ij}(\lambda, \mu)$ to form polynomial invariants similar to those for the Hessian L_{ij}.

DEFINITION 11.

$$\mathcal{I}_n(\lambda, \mu) \stackrel{\text{def}}{=} H_{i_1 i_2}(\lambda, \mu) H_{i_2 i_3}(\lambda, \mu) \cdots H_{i_n i_1}(\lambda, \mu).$$

On expanding this product we find

$$\mathcal{I}_n(\lambda, \mu) = \mu^n I_n + \sum_{k=0}^{n-1} \binom{n}{k} \lambda^{n-k} \mu^k S_0^{n-k-1} S_k. \tag{13}$$

Since $\{\mathcal{I}_1(\lambda, \mu), \ldots, \mathcal{I}_D(\lambda, \mu)\}$ is an irreducible system for the system $\{H_{ij}(\lambda, \mu)\}$ involving only the sets $\{S_0, \ldots, S_{D-1}\}$ and $\{I_1, \ldots, I_D\}$, we conclude that these two sets are *sufficient* for constructing any mixed first- and second-order polynomial invariant. That they are also *necessary* follows by a simple counting argument: there are exactly $2D$ independent invariant degrees of freedom in $\{L_i, L_{ij}\}$ (which is obvious in a coordinate system in which the Hessian is diagonal).

By including the independent, zeroth-order image value L we have finally proved the irreducibility of $\{L, S_0, \ldots, S_{D-1}, I_1, \ldots, I_D\}$ for the case of the 2-jet tensors $\{L, L_i, L_{ij}\}$, which completes the proof of Proposition 2.

Acknowledgments

We are indebted to J. Blom, R. van Maarseveen, and M. van Eert, for their contribution to the software implementation of differential invariants. We also thank A. Salden for his in-depth study of the old mathematical literature on invariants theory, and we thank A. van Doorn for her useful comments on preliminary versions of the manuscript.

This work has been carried out as part of the national priority research program 3D Computer Vision supported by the Netherlands Ministries of Economic Affairs and Education and Science through a SPIN grant and by the industrial companies Philips Medical Systems, KEMA, and Shell Research.

Notes

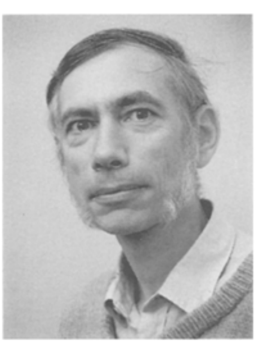
1. We always include the zeroth-order derivative.
2. We sometimes omit the attribute "pseudo" and loosely speak of "invariants" and "tensors." If necessary, the attribute "absolute" is used to exclude pseudo invariants and tensors explicitly.
3. So the (v, w) coordinates serve as Cartesian coordinates of a *tangent space* attached to the given base point, which is defined as the product space of the isophote and gradient tangent spaces at that point. Since we are interested only in local properties defined at the origin of this tangent space, the construction of this Cartesian frame does not serve to derive approximate results but really yields exact results. The price we have to pay for its simplicity is that we are bound to a given base point (the origin of the local frame). If one is interested in multilocal properties, it is inevitable that one either relate the Cartesian frames attached to neighbouring points (by Cartan's method of *moving frame fields*) or introduce *curvilinear coordinates* that parameterise a full spatial neighbourhood instead of our local frame coordinates. Both methods establish a so-called *connection*, i.e., an *orthogonal* and a *metrical connection*, respectively. We will not elaborate on this but will just point out a potential cause of confusion that may arise from the fact that similar notations are encountered in the literature for these totally different kinds of coordinatisations.

References

1. E. Cartan, "Les problèmes d'équivalence," in *Oeuvres Complètes*, vol. 2, pp. 1311-1334, Gauthiers-Villars: Paris, 1952.
2. P.J. Olver, "Invariant theory, equivalence problems and the calculus of variations," in *Invariant Theory and Tableaux*, Springer-Verlag: New York, 1988.
3. H. Weyl, *The Classical Groups, Their Invariants and Representations*, Princeton University Press: Princeton, NJ, 1946.
4. A. Clebsch, *Theorie der Binären Algebraischen Formen*, Verlag von Teubner: Leipzig, 1872.
5. J.A. Dieudoné and J.B. Carrell, *Invariant Theory – Old and New*, Academic Press: New York, 1971.
6. P. Gordan, *Vorlesungen über Invariantentheorie*, Verlag von Teubner: Leipzig, 1885–1887.
7. R. Weitzenböck, *Invariantentheorie*, Noordhoff: Groningen, The Netherlands, 1923.
8. B. Gurevich, *Foundations of the Theory of Algebraic Invariants* Noordhoff, Groningen, 1979.
9. J.H. Grace and A. Young, *Algebra of Invariants*, Chelsea: New York, 1965.
10. I. Schur, *Vorlesungen über Invariantentheorie*, Noordhoff: Groningen, The Netherlands, 1968.
11. D. Hilbert, "Ueber einen algemeinen Gesichtspunkt fur Invarianten-theoretesische Untersuchungen im binaren Formengebiete," *Math. Ann.*, vol. 28, pp. 381–446, 1887.
12. D. Hilbert, "Ueber die Theorie der algebraischen Formen," *Math. Ann.*, vol. 36, pp. 473–534, 1890.
13. F. Klein, "Erlanger Programm," *Math. Ann.*, vol. 43, pp. 63–100, 1893.
14. P. Gordan, "Ueber die Bildung der Resultante zweier Gleichungen," *Math. Ann.*, vol. 50, pp. 355–414, 1871.
15. H. Weber, *Lehrbuch der Algebra*, vol. I-III, Chelsea: New York, 1894.
16. P. Gordan, "Die Discriminate der Form 7. Grades," *Math. Ann.*, vol. 31, pp. 566–600, 1888.
17. M.R. Perrin, "Subdiscriminant," *Mathematique*, vol. 4, pp. 129–167, 1894.
18. H.W. Turnbull, *Theory of Determinants, Matrices, and Invariants*, Dover: New York, 1960.
19. B.L. van der Waerden, *Moderne Algebra*, vol. I–II, Springer-Verlag: Berlin, 1940.
20. J.P.S. Kung and G.C. Rota, "The theory of binary forms," *Bull. Amer. Math. Soc.*, vol. 10, pp. 27–85, 1984.
21. D. Hilbert, "Ueber die vollen Invariantensystemen," *Math. Ann.* vol. 42, pp. 313-373, 1893.
22. D. Stanton, ed., *Invariant Theory and Tableaux The IMA Volumes in Mathematics and Its Applications*, Springer-Verlag: New York, 1988.
23. D.E. Littlewood, "Invariant theory, tensors and group characters," *Philos. Trans. Roy. Soc. London Ser. A.*, vol. 239, no. 807, pp. 305-356, 1944.
24. D. Forsyth, J.L. Mundy, A. Zisserman, C. Coelho, A. Heller, and C. Rothwell, "Invariant descriptors for 3-D object recognition and pose," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 13, pp. 971–991, 1991.
25. L. Van Gool, M.H. Brill, E.B. Barrett, T. Moons, and E. Pauwels, "Semi-differential invariants for nonplanar curves," in *Applications of Invariance in Vision*, J. Mundy and A. Zisserman, eds., pp. 293–309, MIT Press: Cambridge, MA, 1992.
26. L.M.J. Florack, B.M. ter Haar Romeny, J.J. Koenderink, and M.A. Viergever, "General intensity transformations and second order invariants," in *Theory and Applications of Image Analysis, Series in Machine Perception and Artificial Intelligence* vol. 2., P.Johansen and S. Olsen, eds. World Scientific: Singapore, 1992, pp. 22–29.
27. A. Witkin, "Scale space filtering," in *Proc. International Joint Conference on Artificial Intelligence*, Karlsruhe, West Germany, 1983, pp. 1019–1023.
28. J.J. Koenderink, "The structure of images," *Biol. Cybernet.*, vol. 50, pp. 363–370, 1984.
29. T. Lindeberg, "Scale-space for discrete signals," *IEEE Trans. Patt. Anal. Mach. Intell.* vol. 12, pp. 234–245, 1990.
30. L.M.J. Florack, B.M. ter Haar J. Romeny, J.J. Koenderink, and M.A. Viergever, "Scale and the differential structure of images," *Image Vis. Comput.*, vol. 10, pp. 376-388, 1992; also in Computer Vision Research Group, Utrecht, The Netherlands, 3DCV Tech. Report 91–30, 1991.
31. P.J. Olver, *Applications of Lie Groups to Differential Equations, Graduate Texts in Mathematics*, vol. 107, Springer-Verlag: Berlin, 1986.
32. T. Poston and I. Steward, *Catastrophe Theory and Its Applications*, Pitman: London, 1978.
33. J.J. Koenderink and A.J. van Doorn, "Receptive field families," *Biol, Cybernet.*, vol. 63, pp. 291–298, 1990.
34. B.M. ter Haar Romeny, L.M.J. Florack, J.J. Koenderink, and M.A. Viergever, "Scale-space: its natural operators and differential invariants," in *Information Processing in Medical Imaging, Lecture Notes in Computer Science*, vol. 511, A.C.F. Colchester and D.J. Hawkes, eds., Springer-Verlag: Berlin, 1991, pp. 239–255.
35. J. Blom, B.M. ter Haar Romeny, A. Bel, and J.J. Koenderink, "Spatial derivatives and the propagation of noise in gaussian scale-space," *J. Vis. Comm. Image Repr.*, vol. 4, pp. 1-13, March 1993; also in Computer Vision Research Group, Utrecht, The Netherlands, 3DCV Tech. Report 91–03, 1991.
36. J.J. Koenderink, "The brain a geometry engine," *Psychol. Res.*, vol. 52, pp. 122-127, 1990.
37. L. Schwartz, *Théorie des Distributions, Actualités scientifiques et. industrielles*, vol., I, II, Publications de l'Institut de Mathematique de l'Université de Strasbourg, Hermann: Paris, 1950, 1951.
38. L.M.J. Florack, B.M. ter Haar Romeny, J.J. Koenderink, and M.A. Viergever, "Images: regular tempered distributions," to appear.
39. J. Canny, "A computational approach to edge detection," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 8, pp. 679–698, 1986.
40. J.J. Clark, "Authenticating edges produced by zero-crossing algorithms," *IEEE Trans. Patt. Anal. Mach. Intell.* vol. 11, pp. 43–57, 1989.

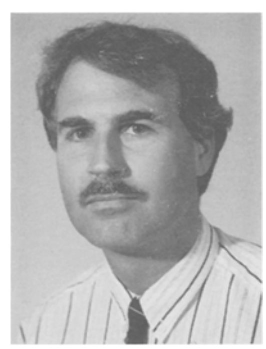
41. J. Blom, *Topological and geometrical aspects of image structure*, Ph.D., thesis, University of Utrecht, The Netherlands, 1992.
42. M.A. Piech, "Decomposing the Laplacian," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 12, pp. 830–831, 1990.
43. A.H. Salden, L.M.J. Florack, and B.M. ter Haar Romeny, "Differential geometric description of 3D scalar images," Computer Vision Research Group, Utrecht, The Netherlands, 3DCV, Tech. Report 91–23, 1991.
44. L.M.J. Florack, B.M. ter Haar Romeny, J.J. Koenderink, and M.A. Viergever, "General intensity transformations," in *Proc. 7th Scandinavian Conference on Image Analysis*, P. Johansen and S. Olsen, eds., Aalborg, Denmark K, August 1991, pp. 338–345; also in Computer Vision Research Group Utrecht, The Netherlands, 3DCV, Tech. Report, 90–20, 1990.
45. P.A. van den Elsen, *Multimodality matching of brain images*, Ph.D. thesis, University of Utrecht, The Netherlands, 1993.
46. J.B.A. Maintz, P.A. van den Elsen, and M.A. Viergever, "Extraction of invariant ridgelike features for CT and MRI brain image matching," in *Proc. Visual Image Processing*, Utrecht, The Netherlands, June, 2–4, 1993.
47. D.F. Lawden, *An Introduction to Tensor Calculus and Relativity*, Spottiswoode Ballantyne, 1962.
48. D.C. Kay, *Tensor Calculus*, Schaum's Outline Series, McGraw-Hill Book Company: New York, 1988.

Luc Florack received his M.Sc. degree in theoretical physics, with a thesis on the quantization of gauge field theories, from the University of Utrecht, The Netherlands, in 1989. He is currently a Ph.D. student in the Computer Vision Research Group, a member of the Utrecht Biophysics Research Institute. His primary research interest in computer vision is the representation of scalar image structure and, in particular, scale-space methods.

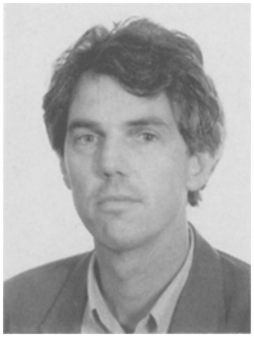


Jan Koenderink received the M.Sc. degree in physics and mathematics in 1967 and the Ph.D. degree in 1972 from the University of Utrecht, The Netherlands. He was an associate professor of experimental psychology at Groningen University until 1974, when he returned to Utrecht, where he presently holds a chair in the department of physics and astronomy. He is currently scientific director of the Utrecht Biophysics Research Institute, in which multidisciplinary work in biology, medicine, physics, and computer science is coordinated. His research interests include optically guided behavior, computational neuroscience, differential geometry, and image processing and interpretation.

Dr. Koenderink received an honorary (D.Sc.) degree in medicine from the University of Leuven and is a fellow of the Royal Netherlands Academy of Arts and Sciences. He participates on the editorial boards of several scientific journals.



Bart M. ter Haar Romeny received an M.Sc. in applied physics from Delft University of Technology in 1978 and a Ph.D. from the University of Utrecht, The Netherlands, in 1983. After being the principal physicist of the Utrecht University Hospital Department of Radiology, he joined the University of Utrecht 3D Computer Vision Research Group as an associate researcher in 1989. His interests are mathematical aspects of front-end vision, particularly, linear and nonlinear scale-space theory, medical computer-vision applications, picture archiving and communication systems, differential geometry and perception, and cross-fertilization among these fields. He is the author of several papers and book chapters on these issues and is involved in (and initiated) a number of international collaborations on these subjects.



Max Viergever received the M.Sc. degree in applied mathematics in 1972 and the D.Sc. degree with a thesis on cochlear mechanics in 1980, both from Delft University of Technology, The Netherlands. From 1972 to 1988 he was assistant professor, and then associate professor of applied mathematics at Delft University. Since 1988 he has been professor of medical image processing and head of the Computer Vision Research Group at the University of Utrecht. He is the author of over 100 scientific papers on biophysics and medical image processing and is the author or editor of nine books. His research interests comprise all aspects of computer vision and image processing, including image reconstruction, compression, multimodality integration, multiresolution segmentation, and volumetric visualization. Dr. Viergever is at present associate editor of *IEEE Transactions on Medical Imaging*.