The Intrinsic Structure of Optic Flow Incorporating Measurement Duality

Luc Florack Wiro Niessen Mads Nielsen

Abstract

The purpose of this report[1] is to define optic flow for scalar and density images without using *a priori* knowledge other than its defining conservation principle, and to incorporate *measurement duality*, notably the scale-space paradigm. It is argued that the design of optic flow based applications may benefit from a *manifest* separation between factual image structure on the one hand, and goal-specific details and hypotheses about image flow formation on the other. The approach is based on a physical symmetry principle known as *gauge invariance*. Data-independent models can be incorporated by means of admissible *gauge conditions*, each of which may single out a distinct solution, but all of which must be compatible with the evidence supported by the image data. The theory is illustrated by examples and verified by simulations, and performance is compared to several techniques reported in the literature.

1 Introduction

The conventional "spacetime" representation of a movie as a sequence of consecutive time frames could be called pseudo-static in the sense that it does not explicitly account for a kinematic relation between local image samples. Such relations naturally arise as a consequence of apparent conservation laws. It is for this reason that the concept of "optic flow" has been introduced by Gibson in the context of optical pilot navigation [26].

Nowadays optic flow has become a familiar, yet still confusing concept in computer vision and image analysis. A meaningful definition depends very much on data-independent models. It is therefore of fundamental importance to study optic flow in the context of structural conventions and data evidence without reference to subjective factors.

The approach adopted in this article is inspired by original work of Arnspang [2, 3, 4, 5], which in turn pursues the classical approach of Horn and Schunck [31]. However, the novelty of this article is threefold:

- *Measurement duality* is taken into account. More specifically we think of optic flow extraction as a measurement process on raw data that cannot be decoupled from the fiducial collection of image processing filters used to probe the data. The principle of duality is explained in Section 2.1.
- We do not incorporate semantics beyond a level needed to define image structure ("proto-semantics"). This gives leeway to bring in prior knowledge and assertions about image formation, task, *et cetera*, which is needed for specific applications. We point out how this can be accomplished in the setting of a so-called *gauge field theory*, the idea of which is outlined in Section 2.2.
- We propose a robust computational recipe, different from existing schemes, to assess the flow up to any desired order of approximation. This is the subject of Section 3.

An implication of duality is that our optic flow field is not of a "speedometer" type (one velocity per base point), but is defined by virtue of a set of fiducial "point operators". In particular we consider the familiar Gaussian scale-space paradigm [35, 39, 84], and show in which precise sense the problem of "deep structure" naturally carries over to optic flow. This problem is induced by the fact that there are

[1] Accepted for publication in the *International Journal of Computer Vision*.

no *a priori* preferred spatial and temporal inner scales. Scale preferences are most sensibly inferred *a posteriori* from image structure given a practical task and a suitable model [55, 56, 58, 64, 65, 66]. In particular, spatial and temporal scales can be adjusted to flow field, and *vice versa*, which seems natural if only for considerations of dimensional analysis: velocity and spacetime scales must be physically dependent. In particular, motion transparency is not precluded from the outset.

For the sake of simplicity we consider images in which grey-value represents a relevant physical parameter associated with some conserved quantity. We distinguish two usual cases, one in which the quantity of interest is a *scalar*, and one in which it represents a *density* [15]. For the sake of definiteness one may think of depth in the case of range imaging, and of proton density cine-MR, respectively. (Caution: "shading" does not fall in either category unless by way of approximation or under suitable conditions.) For reasons of brevity we concentrate on the scalar case, and only briefly touch upon the density case. Section 4 serves as a test and illustration of the theory based on an analytically tractable stimulus. In Section 5 we make a conceptual comparison with similar models proposed in the literature, notably by Otte and Nagel [68] and by Werkhoven and Koenderink [83]. We also describe a quantitative study on a benchmark sequence, and compare performance with results from existing techniques as reported by Barron *et al.* [7]. We summarise results in Section 6 and discuss some aspects relevant for applications and future development.

2 Preliminaries

This section serves to introduce some preliminary concepts that are essential in order to appreciate the precise definition of optic flow, which will be presented in the next section. We also establish the notation.

2.1 The Duality Paradigm

It is impossible to conceive of a grey-value sample without accounting for some kind of "measurement aperture". Thus in order to make inferences one has to rely on mutual interaction between a "naked" source field (a raw image sequence in digital format, say) and a conventionally designed class of probing devices, or filters. This state of affairs is familiar to mathematicians as *duality* [8, 9, 74, 78], and is a public fact to physicists.

Needless to say that duality affects image analysis. It introduces an inevitable bias; we are interested in the source, not in the device. To some extent the bias will be eliminated if we consider a large ensemble of independent devices, or *filter bank*. The latter then becomes the descriptive paradigm for modelling the source of interest [17, 18].

A convenient way to proceed is to define a *linear* filter space, Δ say. The set of all continuous mappings $\Delta \to \mathbb{R}$ is known as its *topological dual*, and is denoted Δ'. Thus instead of representing raw images as functions we may consider them as mappings of localised filters into the real numbers ("local samples").

For general image processing purposes one could take smooth filters of rapid decay, $\mathcal{S}(\mathbb{R}^n)$ (*Schwartz space*) [9, 17, 18, 21, 72]. "Rapid decay" is a condition to enforce filter confinement. Smoothness and rapid decay are strong topological requirements, yet insignificant constraints from a pragmatic point of view. This is so because the raw image data are modelled by the topological dual $\mathcal{S}'(\mathbb{R}^n)$ (*tempered distributions*), which is richer in structure than any of the prototypical Banach spaces $L^p(\mathbb{R}^n)$ usually encountered to model such data ($p \geq 1$, typically $p = 1, 2, \infty$). In particular we decline from assertions about regularity of our data.

A nice property of $\mathcal{S}(\mathbb{R}^n)$ is that it is compatible with image processing demands: if we sequentially filter our source data we obtain a result which can be explained in terms of yet another filter from the same filter class; $\mathcal{S}(\mathbb{R}^n)$ is said to form a (convolution) *algebra*. Thus a sample of a processed image will be a sample according to the original definition (a probe of device space). This is almost never the case with an arbitrarily chosen filter class. A disadvantage is that $\mathcal{S}(\mathbb{R}^n)$ is too large to be implemented. Selecting a finite subset will quite likely cause loss of generality and violation of the consistency requirement (this need not be a problem for *specific* purposes). The paradox can be solved

by postulating a *point operator*, a *positive* filter that generates an *autoconvolution algebra*. There is in fact only one in $\mathcal{S}(\mathbb{R}^n)$, apart from its width and base point, viz. the normalised Gaussian. In this way conventional scale-space theory arises naturally from topological duality.

In general, topological duality admits *well-posed* and *operationally well-defined*[2] differentiation. The trick is "transposition" of operators (possibly ill-posed or even ill-defined) into a suitably designed filter space in which such operators are well-behaved. The idea behind transposition is well-known from linear algebra. Consider a pair of operands taken from two linear spaces, $(u, v) \in U \times V$, say, and suppose one has a recipe for mapping each such pair into a number (a duality principle or a scalar product, for instance). If a linear transformation $S(u) \in U$ has the same effect on the mapping as a linear transformation $T(v) \in V$, in the sense that $(S(u), v)$ produces the same result as $(u, T(v))$ for all (u, v), then T is called the transposed operator associated with S (*vice versa*), notation: $T = S^\dagger$ $(S = T^\dagger)$. For example, if (F, ϕ) is a source-detector pair, and $F[\phi]$ the numerical response from mutual interaction, then by definition $\nabla F[\phi] = F[\nabla^\dagger \phi]$. The crux in this example is that transposition is merely a mental process: the reason why the l.h.s. makes sense is by virtue of the r.h.s. (the operational part!), where differentiation *precedes* discrete sampling. It is easy to see that in case of topological duality we have $\nabla^\dagger = -\nabla$, in other words, that differentiation is anti-symmetric.

Combined with the point concept this leads one to consider the *Gaussian family*, $\mathcal{G}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$, *i.e.* the class of all derivatives of the basic point operator [37, 45]. It is a *complete* family, and thus does not limit potential image analysis. The topological dual $\mathcal{G}'(\mathbb{R}^n)$ is nothing but the well-known *scale-space representation* [35, 84], more precisely, its stratification into local jets of successive orders [23, 27, 43, 48, 49, 69]. It has a straightforward implementation up to some order and within physical scale limits (grain and scope). We will henceforth assume familiarity with the basics of scale-space theory [6, 19, 20, 22, 23, 33, 34, 35, 38, 40, 41, 42, 43, 44, 46, 53, 54, 56, 57, 58, 84]. Especially its interpretation in the context of topological duality will turn out to be crucial for our optic flow definition.

2.2 Aperture Problem and Optic Flow Ambiguity

It is taken for granted that optic flow is a *vector field* [73, 74]. This reflects the desire to link corresponding points—whatever these may be—separated by arbitrarily small temporal intervals. The motivation for this is of course that in the physical world such pointwise connections are actually meaningful; ideally they correspond to particle motion or wave phenomena.

Paradoxically, there is still a lack of public consensus on an unambiguous definition. Broad support exists with regard to the assertion that optic flow is constrained by some *conservation principle*. The classical example is the well-known "Optic Flow Constraint Equation" [2, 3, 4, 5, 7, 31, 32, 60, 61, 68, 71, 75, 76, 79, 83]. If we adopt public consensus as our guideline, then optic flow must be a *homotopy* rather than a vector field; see Figure 1.

The ambiguity of pointwise connections is known as the *aperture problem* [28, 59, 81]. As soon as one attempts to extend the optic flow definition beyond the intrinsically defined homotopy, one has to come up with a *model* in order to "solve the aperture problem". Different models yield different solutions, none of which can be falsified on the basis of data evidence, yet some of which might provide an efficacious solution able to successfully support an application, while others might fail. Thus optic flow disambiguation requires a model as well as a task for its justification, which of course explains the lack of consensus.

A source of confusion about the aperture problem is its association with the existence of "straight edges". This naturally fuels the frequently heard argument that "the aperture problem is a false problem", one that can be overcome as soon as "enough structure" is present in the image brightness, e.g. by taking into account "corner points", or the image's "higher order differential structure" [60, 61, 62, 68, 75, 76, 79, 82, 83]. That this is a misconception follows immediately from the trivial invariance of image structure under isophote automorphisms, making any tangential flow component conceivable (recall Figure 1). Data evidence does not compel us to make any choice whatsoever. Thus from a syntactical point of view, the aperture problem is not a problem, but a (generic) *invariance property*.

[2] A definition is called "operational" if it is algorithmically stated; as such it must only contain references to unambiguous machine concepts.

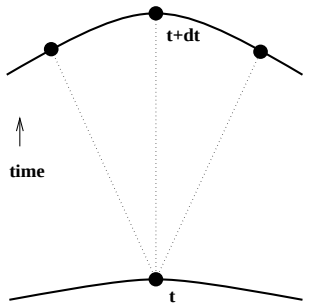

Figure 1: There is equal support from the data for all vectors connecting points on corresponding iso-grey-level contours at two successive moments. The time-parametrised transition from a given isophote to its successor *as a whole* is called a homotopy (the assertion here is that isophote topology does not change). Data evidence does not support pointwise connections, thus optic flow cannot exist as an intrinsic image property. Gauge transformations—explained in the text—are all diffeomorphisms confined to isophotes; under such transformations a given optic flow field is mapped to one that is equivalent in the sense that there is *no observable effect* on the data. The logical consequence is that one cannot get around the aperture problem unless by way of *data-independent* models.

If any, the true problem must be one of semantics. Semantics must be subordinate to, but cannot be deduced from mere structure. Indeed, a closer look into the literature reveals that, in one way or another, *one always brings in semantics* in optic flow disambiguation schemes. Because it is essential to have a handle on semantics for the purpose of validation and, if necessary, revision, it is desirable to *manifestly* separate intrinsic and extrinsic optic flow d.o.f.'s[3], only to combine them in the final stage. By definition, *syntactical optic flow* reflects the intrinsic d.o.f.'s and is *conventionally defined* in terms of the spatiotemporal structure of the image data. Once defined, structure is evidence; it is pointless to make assumptions about something one already knows.

In physics, the segregation into intrinsic and extrinsic d.o.f.'s is manifest in so-called *gauge field theories*. Transformations confined to the subspace of extrinsic d.o.f.'s are called *gauge transformations*. Evidence depends only on the intrinsic d.o.f.'s and is said to be *gauge invariant*. In a gauge field, the extrinsic d.o.f.'s merely enter as disposable variables introduced by an observer to simplify the description of a gauge invariant physical system. The essence of gauge field theories is that the redundancy of auxiliary variables is *implicit* by virtue of a nontrivial gauge invariance of the system at hand. Otherwise it would be senseless to consider them in the first place. Figure 2 illustrates the general idea behind a gauge theory, and the way optic flow fits into such a framework.

Indeed, syntactical optic flow is a gauge field. A "standard" (but in principle *ad hoc*) representation of the factual homotopy is *normal flow*, sometimes called *the* optic flow field, to be distinguished from the physically induced image velocity or *motion field* [29, 30]. A gauge transformation, *in casu* any isophote automorphism, maps one admissible vector field onto another without affecting the data. The extra constraint introduced to single out a unique solution is called a *gauge condition*; see Figure 1. The gauge condition provides a clear entry point for bringing in prior knowledge [29]. Thus the gauge condition expresses the semantics, and the gauge-fixed solution could be called the *semantical optic flow field*, a "meaningful" member of the metamerical class of syntactically equivalent fields. It is important to appreciate that its significance is only relative to the interpretation implied by the gauge condition, and ultimately, relative to the performance of a practical task. For this reason it is equally important to keep syntax and semantics nicely apart until we know exactly what we are after, and even at that stage we may want to re-interpret, change our minds, in other words, alter our gauge in a feedback kind of fashion until we get things to work properly. After all, any *a priori* assertion we make may be wrong.

Considerations for fixing the gauge are basic image formation details, as well as presumptive object attributes such as rigidity or non-elasticity constraints for solid objects [12, 14], incompressibility and continuity conditions for fluids [1, 11], *et cetera*, and of course even the postulate of the very existence of such objects (segmentation and classification). Also, mathematical constraints of smoothness are

[3] d.o.f.('s) = "degree(s) of freedom".

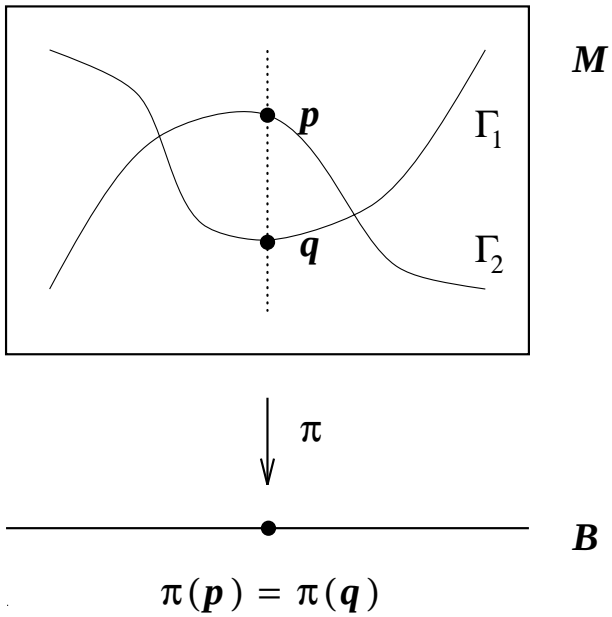

Figure 2: In a gauge theory one considers a manifold $\boldsymbol{M}$ of higher dimensionality than the actual base space of physical observables $\boldsymbol{B}$. The "superfluous" degrees of freedom introduced in this way $(\boldsymbol{p}, \boldsymbol{q}, \ldots)$ effectively cancel out as a result of a symmetry principle postulated for $\boldsymbol{M}$ ("gauge invariance"), *i.e.* there is a stratification into "invariant orbits" that induce identical observations on $\boldsymbol{B}$ (by virtue of the "projection map" π), irrespective of the gauge condition $(, {}_1, , {}_2, \ldots)$ one chooses to impose (as long as it is admissible, *i.e.* transversal to the orbits). The manifold $\boldsymbol{M}$, together with its inherent symmetry, thus merely enters as a "model space" for $\boldsymbol{B}$, in which the manifestation of the latter may be significantly simpler than in any non-gauge model confined to $\boldsymbol{B}$. In the case at hand $\boldsymbol{M}$ will be the vector bundle of n-vectors v representing all possible optic flow vector fields compatible with the image data, while $\boldsymbol{B}$ comprises the actual optic flow evidence in the form of, say, the normal flow field components (more precisely, the 1-dimensional homotopy of iso-grey-level contours). The virtue of a gauge theoretical approach to optic flow lies in the fact that one can model $\boldsymbol{B}$ via a simple, viz. linear model for $\boldsymbol{M}$, at the price of only a mild concession: local redundancy, or gauge invariance. In this case there are $n-1$ gauge degrees of freedom (per base point), which can be fixed by imposing equally many gauge conditions, none of which contradicts data evidence ("tangential flow"). These gauge conditions must reflect *a priori* knowledge or a hypothesis about the physical cause of the induced flow.

often used to fix the gauge. Typically neither one of these provides a globally acceptable constraint; the nature of the gauge depends on the adequacy of local semantics. For example, in the context of machine vision none of the assertions that motion is induced by projection of a shaded, sufficiently smooth, rigid surface patch, *et cetera* [29, 31, 77, 80], is supported by the evidence; at best there is no contradiction. In medical imagery blood flow satisfies physical incompressibility and continuity constraints, whereas bone tissue induces rigid motion, soft tissue deforms nonrigidly, *et cetera*, but one does not know in advance which applies where. The general idea behind such motion constraints—not to be confused with the OFCE—is invariably the same; they express optic flow *coherences* implied by *specific assumptions* about source field coherences, the verification of which must be measured relative to the performance of specific algorithms. In the exploitation of the semantically inspired gauge system, therefore, trial-and-error, feedback, nonlinearities, or whatever we choose to call the successive refinement leading to a consistent gauge, will be rule rather than exception, a process akin to the hermeneutic circle in the transcription of an unknown language (conditionally accept tentative explanations that improve global coherence of the context). The gauge invariant subsystem, on the other hand, is fully determined by a data driven feedforward procedure, and is basically a kinematic reformatting of spatiotemporal image data.

2.3 Computational Problems

There are many possible approaches to optic flow measurement [7, 36], not all of which have been investigated in-depth. The classical approach based on the OFCE defines the optic flow vector field locally by means of a conservation law; invariant grey-values are attributed to points which are dragged along the flow. Its traditional formulation is less than fortunate though, since it bypasses measurement duality. A related computational problem is *ill-posedness*, since the method relies on conventional ill-posed differentiation, and so one has to incorporate a regularisation scheme. Since the outcome of ill-posed problems depends crucially on the details of regularisation, only a thorough motivation of the latter will guarantee that we don't end up with a grab-bag solution.

Although this way of reasoning is in principle legitimate and reflects the usual way of handling the OFCE, we feel that it is somewhat indirect. Since one clearly *has to* regularise ill-posed problems, one might consider recasting the OFCE into a well-posed problem from the very outset. The two conceptual problems, the unavoidable dependency on regularisation method and the problem of ill-posedness, can be played off against each other; reformulating the OFCE in the framework of topological duality solves both problems simultaneously.

2.4 Notation

We will adhere to the following notation. A point in spacetime is indicated by x, its coordinates by x^μ, with $\mu = 0, \ldots, d = n-1$, in which d is the dimension of space (typically 2 or 3), and n that of spacetime. Index values 0 refer to time. Summation convention is implied for pairs of matching indices: $a_\mu b^\mu \equiv a_0 b^0 + \ldots + a_{n-1} b^{n-1}$. Readers not familiar with tensor parlance may choose to ignore the distinction between upper and lower indices at the expense of one mild concession: that all coordinates are taken relative to a *Cartesian frame*. In integral expressions we use a prefix notation for the measure dx, so that it precedes everything in the integrand that depends on x. The symbol $\mathbb{Z}_0^+$ is used to denote all natural numbers including zero.

The source field of interest is modelled either by a scalar function f (any reasonable model for the raw pixel data will do), or by a corresponding linear continuous functional $F[\phi]$, if we want to conceive of it as a mapping of filters $\phi \in \mathcal{G}(\mathbb{R}^n)$ into the reals. In other words, $F \in \mathcal{G}'(\mathbb{R}^n)$. These two representations are interchangeable, since F is fully determined by f: $F[\phi] = \int dx\, f(x)\phi(x)$ (the so-called *Riesz representation theorem*).

Furthermore, we will reserve the symbol ϕ for the zeroth order Gaussian point operator, and write $\phi_{\mu_1 \ldots \mu_k}$, or $\partial_{\mu_1 \ldots \mu_k}\phi$, for its k-th order derivative w.r.t. $x^{\mu_1}, \ldots, x^{\mu_k}$ (covariant derivatives are understood when using non-Cartesian coordinates). Without loss of generality we will assume that our base point of interest corresponds to the origin of our coordinate system, and for simplicity of notation consider unit

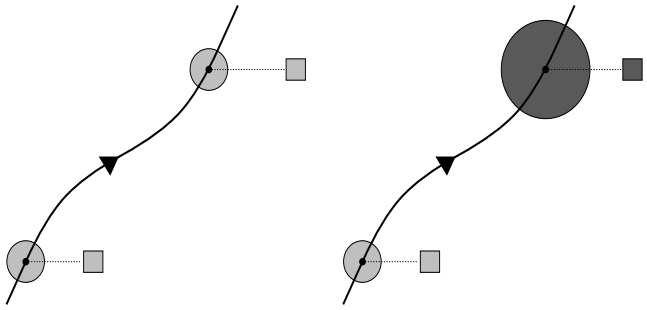

Figure 3: In a *scalar* interpretation of the OFCE it is asserted that *points*, or *fixed-scale* samples taken at these points, lying on one flow line, share the same grey-value attribute. In the *density* case, grey-values are attributed to *volume elements* which are themselves susceptible to the flow. In this sketch, the shaded patches represent corresponding volume elements, the size of which is seen to increase in the density case due to the divergence component of the flow field; conservation entails that their volumetrically integrated grey-values be the same provided the integration aperture is consistently transformed by the flow. White means dense, black means void, background colour has no meaning. The boxes symbolise pixel values.

spatial and temporal inner scales. It is straightforward to make spatial and temporal scale parameters explicit if desired.

An important tool in models of flow and conservation is the so-called *Lie derivative*. Lie derivatives capture variations of spacetime quantities along the integral flow of some vector field. To take a Lie derivative, one therefore needs to know this vector field. Actually, we shall only consider the 1-st order Lie derivative of an image, which will give us a *linear* model of optic flow; this is *not* a restriction and should not be confused with the spatiotemporal differential order of the flow field one might be interested in. We have no reason to impose any *a priori* restrictions on this. The vector field will, of course, be the optic flow field, but note that it is a vector in spatiotemporal sense, including a not necessarily trivial temporal component. We will denote the *infinite resolution*[4] field by $v^\mu \equiv (v^0; \vec{v})$.

The Lie derivative of a scalar function f w.r.t. a vector field v^μ is given by the directional derivative $\mathcal{L}_v f = \partial_\mu f \, v^\mu$. Similarly, if f represents a density, we have $\mathcal{L}_v f = \partial_\mu (f \, v^\mu)$ (Figure 3). However, like any classical derivative expressions like these are ill-posed, and, besides, require f to be differentiable, which is an operationally void constraint for discrete data. Well-posed differentiation is obtained by identifying f with its scale-space representation $F \in \mathcal{G}'(\mathbb{R}^n)$; this brings the filter paradigm into focus.

Having established notations and conventions, we now turn to the theory.

3 Theory

The following assumption will be made in order to connect to the usual optic flow terminology.

Assumption 1 (Temporal Gauge: Conservation of Topological Detail)

$$\forall x \in \mathbb{R}^n : \quad v^0(x) = 1 \, .$$

This is a locally weak, but globally not necessarily realistic assumption, stating that the flow is everywhere nonvanishing and transversal to constant-time frames, in other words, that the topological properties of the isophote topography are preserved over time despite deformations. It therefore expresses *conservation of topological detail*[5]. Thus Assumption 1 is an instance of a gauge condition enforced on the basis of an *a priori* model (Figure 4).

We shall be needing a *formal expansion* of v^μ near the origin, truncated at some arbitrary order M, v^μ_M say. This is a polynomial intended to capture a finite number of local d.o.f.'s of the vector field.

[4] Infinite resolution is of course merely hypothetical, but serves to define the actual multiresolution field via an integral formula similar to the Riesz representation form for the raw image data f. We return to this later.

[5] Conservation—as opposed to invariance—presupposes temporal causality.

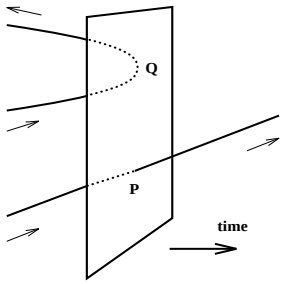

Figure 4: A transversal and a non-transversal flow-line (through **P** and **Q**, respectively). The latter one is excluded by transversality. The upper branch of that curve has an anti-causal orientation. It can be given a causal interpretation simply by reversing the upper arrow; in that case, however, **Q** becomes an annihilation point, thus violating the assumption of "conservation of topological detail". Note that if the *spatial* optic flow field vanishes, the flow-lines will be parallel to the time axis.

Definition 1 (M-th Order Formal Expansion) *The formal expansion of order M of the vector field v^μ, denoted v^μ_M, is an M-th order polynomial*

$$v^\mu_M(x) = \sum_{l=0}^{M} \frac{1}{l!} v^\mu_{M;\rho_1 \dots \rho_l} \, x^{\rho_1} \dots x^{\rho_l} \, ,$$

the coefficients of which may depend on M.

The finite set of coefficients $v^\mu_{M;\rho_1 \dots \rho_l}$ (to be defined later on so as to *approximate* the optic flow field's derivatives $\partial_{\rho_1 \dots \rho_l} v^\mu(x=0)$ at the origin), corresponds to the d.o.f.'s we shall try to solve for. Note that v^μ_M is *not* required to be the M-th order Taylor polynomial of v^μ. The idea of approximation is merely a deviation of order $\mathcal{O}(\|x\|^{M+1})$.

The Lie derivative of the image with respect to the optic flow vector can be made well-posed by formulating it as a tempered distribution.

Proposition 1 (Lie Derivative of a Raw Image) *The Lie derivative of a raw image $F \in \mathcal{G}'(\mathbb{R}^n)$ is defined by virtue of its dual; i.e. for every $\phi \in \mathcal{G}(\mathbb{R}^n)$ we have*

$$\mathcal{L}_v \, F[\phi] \stackrel{\text{def}}{=} F[\mathcal{L}^\dagger_v \, \phi] \, ,$$

in which $\mathcal{L}^\dagger_v \, \phi = -\partial_\mu \, (\phi \, v^\mu)$, if the raw image is a scalar, and $\mathcal{L}^\dagger_v \, \phi = -\partial_\mu \phi \, v^\mu$, if it is a density.

In the subspace of distributions in which one can take classical Lie derivatives of the raw image f as explained in the previous section, one can easily prove Proposition 1 using the integral expression for $F[\phi]$ and subjecting it to a partial integration (*voilà*, the minus sign). Beyond that subspace it is a matter of definition.

Note that if the raw image represents a scalar, then the filter is a density (*n-form* in geometric jargon). Otherwise, if the raw image is a density, we must interpret the filter as a scalar with regard to Lie derivation. This is the plain consequence of duality. For example, in the "generalised motion constraint equation" Schunck considers the flow of a "naked density" field [71]. This result does not account for the essential role of duality. Transposed to filter space we obtain a *scalar* transformation of the filters (the second case in Proposition 1).

Our goal is to reconcile syntactical optic flow with the duality paradigm, notably scale-space theory, and to derive an operational scheme for its computation. We henceforth assume that the source is a scalar; the density case is briefly outlined, but the details are left to the reader.

In the spirit of the traditional approach, we define gauge invariant optic flow as any vector field that preserves $F[\phi]$:

Definition 2 (Syntactical Optic Flow: Global Definition) *See Proposition 1. The syntactical optic flow field is defined as the equivalence class of ∞-resolution vector fields v^μ that satisfy the OFCE*

$$\mathcal{L}_v\, F[\phi] = 0$$

at every base point in spacetime. In integral form this reads:

$$\int dx\, f(x)\, \partial_\mu \left(\phi(x)\, v^\mu(x)\right) = 0\,,$$

assuming the source to be a scalar.

In this way the vector field v^μ and the scalar source data f are coupled in a natural way.

Disambiguation requires us to decide on the gauge. In the absence of *a priori* knowledge we may take—*only for the sake of presentation*—the "standard gauge" that nullifies tangential flow, and in addition asserts that topological detail is conserved.

Definition 3 (Standard Gauge: Normal Flow) *Cf. Definition 2. Normal flow is defined as optic flow subject to the standard gauge*

$$\mathcal{L}_{*v}\, F[\phi] = 0\,,$$

in addition to the temporal gauge of Assumption 1. Here, $*v$ represents the spatial dual of v: $*v_\mu v^\mu \equiv 0$ and $*v^0 \equiv 0$.

Note that there are typically $d-1$ such independent equations, and 1 more for the temporal gauge, per base point and per level of resolution. Altogether we end up with $n = d+1$ field equations for the components of the standard solution, as it should. Note that in practical cases only the gauge equations are replaced. The gauge invariant equation of syntactical optic flow is always the same; image modality and specific task are immaterial. The problem of optic flow extraction (*operational* definition), in whatever meaningful gauge, is to unconfound source data and vector field from the resulting system of equations.

Although formulated for one sample point only, Definition 2 as well as Definition 3 should be understood as *globally valid* identities (as is, trivially, Assumption 1); they remain true if we extend the model from local samples to global images (sets of local samples endowed with a spacetime topology[6] [17, 18]). This can be expressed in terms of a countable set of *local* constraints by stating that all spatiotemporal derivatives at the point of interest vanish as well (caution: Lie derivatives and ordinary derivatives do not commute).

Definition 4 (Syntactical Optic Flow: Local Definition) *Recall Definition 2. At fixed spacetime base point the syntactical optic flow field is equivalently defined as the equivalence class of ∞-resolution vector fields v^μ that satisfy the linear system*

$$\partial_{\mu_1 \ldots \mu_k} \mathcal{L}_v\, F[\phi] = 0 \quad \text{forall} \quad k \in \mathbb{Z}_0^+\,,$$

in which the spatiotemporal derivatives of $\mathcal{L}_v\, F[\phi]$ are defined as usual:

$$\partial_{\mu_1 \ldots \mu_k} \mathcal{L}_v\, F[\phi] = (-1)^k \int dx\, f(x)\, \mathcal{L}_v^\dagger\, \phi_{\mu_1 \ldots \mu_k}(x)\,,$$

with $\mathcal{L}_v^\dagger\, \phi_{\mu_1 \ldots \mu_k} = -\partial_\mu \left(\phi_{\mu_1 \ldots \mu_k}\, v^\mu\right)$.

(The case for a density image should be clear: no derivative on v^μ.) Note that in the static case, *i.e.* if the source is time-independent, a standard solution is $v^\mu = (1; \vec{0})$. In the general case we need an operational method to solve for the optic flow field. In order to accomplish this, the idea pursued below will be to get the field v^μ, defined implicitly by Definition 2, "out of the integral".

In principle, the optic flow field v^μ contains an infinite number of d.o.f.'s. This is inconvenient. Moreover, most of them are irrelevant or computationally inaccessible anyway. However, depending on

[6] Assuming homogeneity one may replace local samples $F[\phi]$ by function valued distributions $F \star \phi\,(x) = \int d\xi\, f(x+\xi)\phi(\xi)$.

one's task, a 0-th order approximation is usually too restrictive. For example, in the case of real-world movies, 1-st order properties of the vector field may reveal relevant information such as qualitative shape properties, surface slant [47], and time-to-collision [51, 52]. Unlike 1-st order, 2-nd order is quantitatively related to intrinsic surface properties of an object [50]. There is no *a priori* limit to the highest order that is still accessible and significant; this depends very much on matters such as image quality (noise and sampling characteristics), resolution of interest, *et cetera*. But independent of such considerations of *differential order*, it will be argued below that we must consider an *approximation order* which is at least equal to, but may well exceed the differential order of interest, in order to provide a viable approximation. In fact, even for zeroth differential order, there is no *a priori* upper limit for an acceptable approximation order $M \geq 0$, though if image quality permits, there will typically be only a set of measure zero at which such an M exceeds computationally realistic limits (flow discontinuities; we return to this briefly in Section 6).

So let us consider the M-th order case. Replacing v^μ by v^μ_M according to Definition 1 yields the following.

Result 1 (M-th Order Optic Flow Approximation) *See Definition 1 and Definition 4. Using $\mathcal{L}_{v_M}$ instead of $\mathcal{L}_v$ we have*

$$\partial_{\mu_1 \ldots \mu_k} \mathcal{L}_{v_M} F[\phi] = -\sum_{l=0}^{M} v^\mu_{M;\rho_1 \ldots \rho_l} \int dx\, f(x)\, \partial_\mu \Phi^{\rho_1 \ldots \rho_l}_{\mu_1 \ldots \mu_k}(x) = 0 \quad \textit{for all } k = 0, \ldots, M \,,$$

in which the effective filters $\Phi^{\rho_1 \ldots \rho_l}_{\mu_1 \ldots \mu_k}$ are given by

$$\Phi^{\rho_1 \ldots \rho_l}_{\mu_1 \ldots \mu_k}(x) = \frac{(-1)^k}{l!}\, \phi_{\mu_1 \ldots \mu_k}(x)\, x^{\rho_1} \ldots x^{\rho_l} \,.$$

Since the Gaussian family $(-1)^k \phi_{\mu_1 \ldots \mu_k}$ $(k \in \mathbb{Z}^+_0)$ is complete, the set of filters $\Phi^{\rho_1 \ldots \rho_l}_{\mu_1 \ldots \mu_k}$ $(k, l \in \mathbb{Z}^+_0)$ is apparently *overcomplete*. Hence they can all be expressed in terms of pure Gaussian derivative filters (the case $l = 0$). For technical details see Appendix A.

Despite the complex appearance of Result 1, it is a straightforward linear system, the coefficients of which are well-posed scale-space derivatives, which can in principle be inverted *analytically* once and for all (for each M). It should also be noted that this system differs in an essential way from similar linear systems in optic flow parameters proposed in the literature: the size of the system is related to the order of approximation in a crucial way explained below, and the precise structure of the coefficient matrix is the consequence of duality plus explicit exploitation of the scale-space paradigm. We return to these observations in Section 4.

It is important to note the restriction on the admissible spatiotemporal orders k. Two essentially equivalent motivations can be given to show that k cannot exceed M.

- Our approximation scheme boils down to replacing the Lie-derivative $\mathcal{L}_v f = \partial_\mu f\, v^\mu$ by $\mathcal{L}_{v_M} f = \partial_\mu f\, v^\mu_M = \mathcal{L}_v f + \mathcal{O}(\|x\|^{M+1})$. Thus we can consider only spacetime derivatives of order $k \leq M$ if we want to keep the unknown remainder small near the point of expansion. Only then can we expect the coefficients of v^μ_M to be close to the Taylor coefficients of v^μ: a k-th order Taylor coefficient of v^μ equals the corresponding coefficient of v^μ_M up to $\mathcal{O}(\|x\|^{M-k+1})$.

- Alternatively, we have to take care not to introduce spurious d.o.f.'s by truncation; the assumption that the Taylor tail of v^μ vanishes identically—a hypothesis not evident from the data—boils down to a gauge condition! We should try to maintain gauge invariance for our *approximated* optic flow field v^μ_M. Allowing arbitrary orders of differentiation in Result 1 would certainly break this invariance in the generic case (*i.e.* the usual case when $v^\mu \neq v^\mu_M$), and may even yield an *inconsistent* system[7]! This is another reason why one needs to limit the highest order to $k = M$. Indeed, it can be shown that the resulting linear equations are generically independent and indeed gauge invariant *to the same extent as the exact system* [24]. Thus the M-th order approximation does *not* affect the intrinsic ambiguity of optic flow at all. Although invariance now relates to *approximations* of isophote automorphisms, no artificial constraints have sneaked in.

[7] Which of course can always be "solved" in least squares sense...

Unfortunately, there are several instances in the literature in which a unique optic flow field is singled out by specific combinations of truncation and differentiation. What really happens is that these very combinations implicitly fix the gauge. One cannot expect this to be a proper way of gauge fixing, except for coincidental cases or by virtue of the validity of unspecified hypotheses. Such methods transcode the aperture problem into optic flow ambiguity; one establishes an instance of an optic flow field, the semantical content of which is not transparent and consequently hard to verify or falsify. Gauge fixing, or optic flow disambiguation, is really a matter of data-extrinsic physical considerations independent of the image, the OFCE, or anything derived from these[8].

An illustrating example—although not accounting for duality—has been given by Amini [1] in the context of 2D X-ray projection imaging, who argues on the basis of a physical model of incompressibility of blood, that the projected flow must be divergence-free; in the application at hand this suffices to fix the gauge (but only for blood flow; clearly one needs different models for different tissues). Devlaminck and Dubus [11] consider density images of deformable media, for which they propose to use the small displacement theory of elasticity in order to establish a physically meaningful gauge in a Tikhonov regularisation framework (again, the gauge entails hypotheses about local tissue characteristics, which calls for a consistency check).

As opposed to conventional schemes based on M-fold implicit differentiation of the OFCE, *every* k-th order subset of Result 1 contains M-th order components of the approximated optic flow field. It is important to keep in mind that the coefficients $v^\mu_{M;\rho_1\dots\rho_l}$ in v^μ_M depend on the order M of approximation. In other words, the polynomial approximation v^μ_{M+1} is a refinement of v^μ_M in the sense that *all* coefficients are refined. Hence, it is *not* the Taylor polynomial of v^μ. Only in the limiting case we have $\lim_{M\to\infty} v^\mu_{M;\rho_1\dots\rho_l} = \partial_{\rho_1\dots\rho_l} v^\mu(x=0)$, so that $v^\mu_\infty = v^\mu$, subject to its original, gauge invariant definition. This brings us to another misconception one frequently encounters in the literature on the OFCE; since a polynomial approximation is typically not a truncated Taylor expansion, one cannot expect *any* choice of M to provide an accurate determination of optic flow derivatives! The principle of refinement is essential.

At this point we have an operational definition of optic flow, built on the foundations of scale-space theory. In order to make the connection transparent we can relate it to diffusion.

Definition 5 (Optic Flow as a Vector-Valued Distribution) *Let the source current J^μ be defined as the linear, vector-valued distribution with Riesz representation $j^\mu = v^\mu f$, in which v^μ is the ∞-resolution optic flow field associated with the raw image f according to previous definitions:*

$$J^\mu[\phi] \stackrel{\text{def}}{=} \int dx\, j^\mu(x)\, \phi(x)\,.$$

Then we can define a corresponding nonlinear, vector-valued distribution V^μ, as follows:

$$V^\mu[\phi] = \frac{J^\mu[\phi]}{F[\phi]} = \frac{\int dx\, j^\mu(x)\, \phi(x)}{\int dx\, f(x)\, \phi(x)}\,.$$

The distribution V^μ is well-defined as long as $F[\phi] \neq 0$.

Thus although the ∞-resolution velocity field v^μ (or v^μ_M in practice) provides us with a speedometer type of measurement (one velocity per base point), the necessity of defining a filter space will always give rise to multiple, yet mutually consistent interpretations (*e.g.* motion transparency may be accounted for). In particular, if ϕ is a 2-parameter spatiotemporal Gaussian, then $V^\mu[\phi]$ defines the multiscale optic flow field induced by f, with v^μ as the hypothetical limit of zero spatial and temporal scale. In that case $J^\mu[\phi]$ (*not* $V^\mu[\phi]$) satisfies the same diffusion equation as $F[\phi]$.

Result 2 (Optic Flow Scale-Space) *See Definition 5 for notation.*

$$\left\{ \begin{array}{lcl} \partial_s J^\mu & = & \frac{1}{2}\Delta J^\mu \\ \lim_{s\downarrow 0} J^\mu & = & j^\mu\,, \end{array} \right.$$

[8] In his original work, Arnspang treats the gauge condition as an explicit *mathematical* constraint.

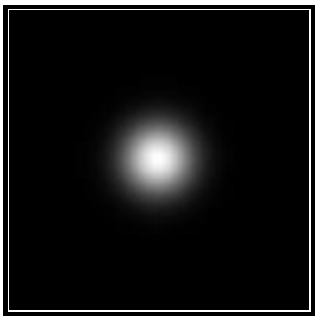

Figure 5: First frame of the optic flow test sequence, corresponding to $t = 0$.

in which the spatiotemporal Laplacean (Laplace-Beltrami operator) is defined as usual, but with respect to pseudo-isotropic coordinates $x^\mu = (ct; \vec{x})$*, and in which the evolution parameter* s *is related to the two independent scale parameters* σ *(isotropic space) and* τ *(time) according to* $s = \sigma^2 = c^2\tau^2$ *for a suitable value of velocity parameter*[9] c*.*

Finally note that the temporal gauge of Assumption 1 is respected, regardless of filter choice, that the linear functional $J^\mu[\phi]$ is always well-defined ("linear flow"), as opposed to the nonlinear functional $V^\mu[\phi]$ ("geometric flow"), and that for $\mu = 0$ Result 2 reduces to the familiar scale-space diffusion equation induced by the raw image f. This shows in which precise way optic flow inherits its "deep structure" from that of its generating scalar source field.

4 Simulation and Verification

In order to test the theory, and at the same time illustrate how one could proceed in specific applications, we define two analytically tractable (2+1)-dimensional stimuli, one simulating a density, the other a scalar field. We then derive an exact, closed-form expression for normal flow in the usual temporal gauge, exploiting the scalar paradigm for *both* cases. This is done for the purpose of analysis and empirical verification, although the construction suggests that in the density case one would probably want to use the density paradigm instead. Also, simulation has the advantage that evaluation will not be hampered by uncertain factors such as reconstruction artifacts, complications that often arise in real image data (and have to be handled on the basis of modality specific models).

Noise generally causes any idealisation to be violated, and so we study significant noise perturbations as well. There ought to be no problem according to theoretical prediction: away from isolated singularities (explained as an artifact of conventional optic flow gauge), everything depends continuously on the noise.

The raw images consist of oscillating Gaussian blobs (Figure 5) that behave as a density and as a scalar, respectively, and the tests are run on discretised floating-point representations of these, as well as on instances perturbed by 50% multiplicative, pixel-uncorrelated Gaussian noise.

4.1 Density Gaussian

Consider the following stimulus definition in a Cartesian coordinate system:

$$F(x, y, t) = \frac{1}{4\pi s(t)} e^{-\frac{x^2+y^2}{4s(t)}} , \tag{1}$$

with the following choice of kinematics:

$$s(t) = A + B\sin(\frac{2\pi t}{T}) \qquad (A > B > 0) . \tag{2}$$

[9] One can opt to discard the formal parameter c and write two diffusion equations for spatial and temporal parts separately.

Conservation for this stimulus follows from

$$\int dx dy\, F(x, y, t) = 1 \quad \forall t\,. \tag{3}$$

In other words, each time-slice contains the same amount of "mass". The Lie-derivative w.r.t. the vector field $\underline{v} = (v^t; v^x, v^y)$ is

$$\mathcal{L}_{\underline{v}} F = F_x v^x + F_y v^y + F_t v^t \stackrel{\text{def}}{=} 0\,, \tag{4}$$

in which a subscript denotes a partial derivative. This result holds globally. The temporal component can locally be gauged to unity as usual,

$$v^t \stackrel{\text{def}}{=} 1\,, \tag{5}$$

provided topological detail is conserved. We will henceforth write (u, v) instead of (v^x, v^y). A straightforward computation yields the following linear equation:

$$xu + yv = \left(\frac{r^2}{2s(t)} - 2\right) \dot{s}(t) \stackrel{\text{def}}{=} \alpha(r, t)\,. \tag{6}$$

Normal flow can be obtained by imposing an additional spatial gauge, which looks identical to Equation 4 except for a replacement of $\underline{v} = (1; u, v)$ by its spatial dual $*\underline{v} = (0; -v, u)$. This yields

$$xv - yu = 0\,. \tag{7}$$

In other words: the spatial velocity vector $\vec{v}$ is everywhere parallel to $\vec{x} = (x, y)$, which is also obvious from considerations of symmetry. Solving Equation 6 subject to the gauges 5 and 7 yields the following solution:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{\alpha(r, t)}{r^2} \begin{pmatrix} x \\ y \end{pmatrix}\,. \tag{8}$$

The origin is problematic in the gauge 5 even though the example is not at all pathological[10]. Note also that $\vec{v} = \vec{0}$ on the oscillating circle defined by $r^2 = 4s(t)$ (F_t vanishes identically), and that the direction of the optical flow vector flips across this circle (which, by the way, runs counter to our visual percept, which gives the impression of alternating expansions and contractions). The situation at time $t \approx 0$ is as follows:

- within the zero-flow radius, $\vec{v}$ points inward, *i.e.* towards the centre of the blob,
- outside the zero-flow radius, $\vec{v}$ points outward, and
- the symmetry centre is a singularity.

The singularity arises as an artifact of the scalar flow paradigm, for which the imposed gauge conditions are apparently too strong. In particular the temporal gauge is not compatible with the topological transitions that occur at the centre, where isophotes continuously pop up or disappear. In reality there ought to be no fundamental problem, of course; the reader may verify that the singularity vanishes when using the density paradigm subject to a natural gauge instead, and calculating physically meaningful quantities, such as the "mass flux" through a sphere containing the singularity. For example, a radial field $\vec{v}(r) \propto \vec{r}/r^2$ in 3D will yield a constant flux (by the divergence theorem).

4.2 Scalar Gaussian

We now consider the scalar case. We synthesise a raw image of the form

$$G(x, y, t) = e^{-\frac{x^2+y^2}{4s(t)}}\,. \tag{9}$$

[10] Note that the singularity cannot be "blurred out"; blurring amounts to an offset in $s(t)$.

For mathematical convenience we take

$$\sigma(t) = A + B\sin(\frac{2\pi t}{T}) \qquad (A > B > 0)\,. \tag{10}$$

Equation 3 does not hold. The Lie-derivative w.r.t. the optical flow vector field is again given by Equation 4, F replaced by G. Again we assume the temporal gauge to hold: Equation 5. Unlike previously, there ought to be no problem with this, since the scalar flow paradigm is the appropriate one to use for this stimulus. We get the following linear equation:

$$xu + yv = \frac{r^2}{2s(t)}\dot{s}(t) \stackrel{\text{def}}{=} \beta(r,t)\,. \tag{11}$$

Normal flow can be obtained as in Equations 5 and 7. With Equation 11 this yields the following solution:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{\beta(r,t)}{r^2}\begin{pmatrix} x \\ y \end{pmatrix}\,. \tag{12}$$

The solution differs qualitatively from Equation 8. First of all, $\beta(r,t)$ has a global sign, equal to that of $\dot{s}(t)$; there is no zero-flow spatial circle across which the flow inverts (in agreement with perceptual flow). Secondly, we indeed encounter no singularity, whereas previously we had

$$\|\vec{v}(r)\| \sim 2\frac{|\dot{s}|}{r} \qquad (r \downarrow 0)\,. \tag{13}$$

4.3 Numerical Test

The frame of interest is the first one, corresponding to $t = 0$. The spatial symmetry centre is taken as the origin $(x, y) = (0, 0)$. Parameter values are as follows. For Equations 1–2 we take $A = 52$, $B = 20$, $T = 16$. For Equations 9—10 we choose $A = 8$, $B = 4$ and $T = 16$. We select the same spatial and temporal scale parameters in both cases: $\sigma = 2$ and $\tau = 1$ (all values relative to pixel scale or frame interval). Both images are of size $16 \times 128 \times 128$, so that they accommodate full periods, and are computed with 4 bytes-per-pixel floating point precision.

We impose the normal flow field equations

$$\begin{cases} \mathcal{L}_{\underline{v}} F[\phi] &= 0 \\ \mathcal{L}_{*\underline{v}} F[\phi] &= 0\,, \end{cases}$$

for $\underline{v} = (1; u, v)$, whence $*\underline{v} = (0, -v, u)$, using the analytical, zeroth or first order approximating schemes (we take Assumption 1 for granted here). The analytical study has been discussed in-depth in the previous sections. We henceforth refrain from mentioning the filter ϕ and its associated scale parameters for the sake of notational simplicity:

$$L_{\mu_1 \ldots \mu_k} \stackrel{\text{def}}{=} (-1)^k\, F[\phi_{\mu_1 \ldots \mu_k}]\,,$$

in which each index μ_i denotes a *scaled* partial derivative, e.g. $L_{xtt} = -F[\phi_{xtt}] = -\sigma\,\tau^2\, F[\partial_{xtt}\phi]$. It is important for the discussion that follows to distinguish between the explicit scaling factors σ and τ for the amplitudes, and the inner scale parameters of the Gaussian derivative filters.

The $k = 0$ approximation is based on a zeroth order polynomial vector field

$$u_0(t; x, y) = u\,, \qquad v_0(t; x, y) = v\,,$$

and boils down to the inversion of

$$\begin{cases} uL_x + vL_y &= -L_t \\ -vL_x + uL_y &= 0\,, \end{cases}$$

which is the same as the traditional OFCE for normal flow. For $k = 1$ we have

$$u_1(t; x, y) = u + u_t t + u_x x + u_y y\,, \qquad v_1(t; x, y) = v + v_t t + v_x x + v_y y\,,$$

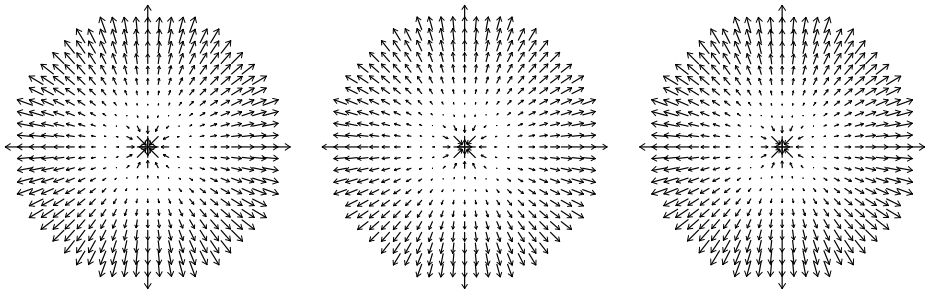

Figure 6: **Left**: Analytically determined velocity field for the density stimulus of Formula 1. **Middle & Right**: Velocity field obtained in the zeroth and first order approximation ($\sigma_s = 2; \sigma_t = 1$).

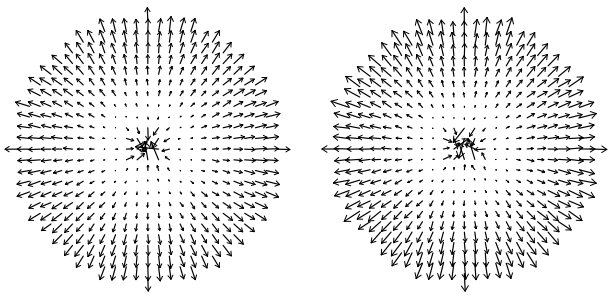

Figure 7: Velocity field obtained for the stimulus of Formula 1. 50% Multiplicative Gaussian noise has been added to the original data. **Left**: zeroth order approximation. **Right**: first order approximation.

and obtain instead

$$\left\{\begin{array}{lcl}
uL_x + vL_y + u_tL_{xt} + v_tL_{yt} + u_xL_{xx} + v_xL_{xy} + u_yL_{xy} + v_yL_{yy} & = & -L_t \\
uL_{xt} + vL_{yt} + u_t(L_x + L_{xtt}) + v_t(L_y + L_{ytt}) + u_xL_{xxt} + v_xL_{xyt} + u_yL_{xyt} + v_yL_{yyt} & = & -L_{tt} \\
uL_{xx} + vL_{xy} + u_tL_{xxt} + v_tL_{xyt} + u_x(L_x + L_{xxx}) + v_x(L_y + L_{xxy}) + u_yL_{xxy} + v_yL_{xyy} & = & -L_{xt} \\
uL_{xy} + vL_{yy} + u_tL_{xyt} + v_tL_{yyt} + u_xL_{xxy} + v_xL_{xyy} + u_y(L_x + L_{xyy}) + v_y(L_y + L_{yyy}) + & = & -L_{yt} \\
uL_y - vL_x + u_tL_{yt} - v_tL_{xt} + u_xL_{xy} - v_xL_{xx} + u_yL_{yy} - v_yL_{xy} & = & 0 \\
uL_{yt} - vL_{xt} + u_t(L_y + L_{ytt}) - v_t(L_x + L_{xtt}) + u_xL_{xyt} - v_xL_{xxt} + u_yL_{yyt} - v_yL_{xyt} & = & 0 \\
uL_{xy} - vL_{xx} + u_tL_{xyt} - v_tL_{xxt} + u_x(L_y + L_{xxy}) - v_x(L_x + L_{xxx}) + u_yL_{xyy} - v_yL_{xxy} & = & 0 \\
uL_{yy} - vL_{xy} + u_tL_{yyt} - v_tL_{xyt} + u_xL_{xyy} - v_xL_{xxy} + u_y(L_y + L_{yyy}) - v_y(L_x + L_{xyy}) & = & 0\,.
\end{array}\right.$$

The lowest order system has $1+1$ equations in 2 unknowns, u, v, and is determined in terms of the image's first order derivatives ($1+2$ equations in 3 unknowns if temporal gauge is made explicit). The first order system comprises $4+4$ equations in $2+6$ unknowns, $u, v, u_x, u_y, u_t, v_x, v_y, v_t$, and requires derivatives of orders $1, 2, 3$ (4+8 equations in 3+9 unknowns, respectively). Recall that these parameters are not the flow field's partial derivatives; for example, the parameters u, v arising from the latter system are to be distinguished from, and generally refine those of the former (the order tag has been left out for notational simplicity).

Both systems have been solved numerically by pixel-wise "LU decomposition" as described in Numerical Recipes [70, section 2.3]. (For a global gauge a significant gain of speed may be accomplished by analytical inversion prior to computation.) The coefficients—in the spatial domain: convolutions of the raw image with Gaussian derivative filters—have been computed using FFT in the straightforward way. Figures 6–17 show the results (vectors are suitably scaled for the sake of visualisation).

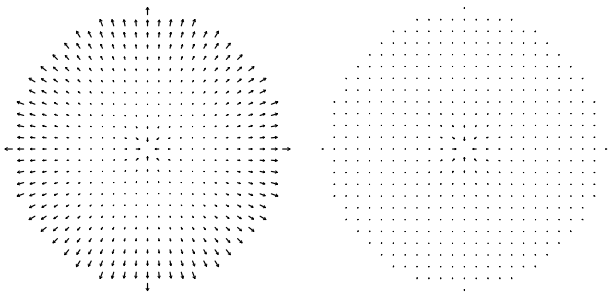

Figure 8: Error in velocity field obtained for the density stimulus of Formula 1 in the noise free case. **Left**: zeroth order approximation. **Right**: first order approximation.

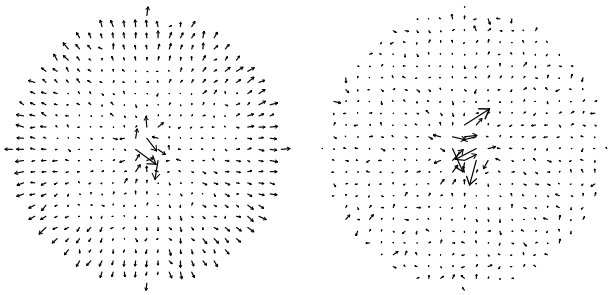

Figure 9: Error in velocity field obtained for the density stimulus of Formula 1 after the original image has been corrupted with 50% multiplicative Gaussian noise. **Left**: zeroth order approximation. **Right**: first order approximation.

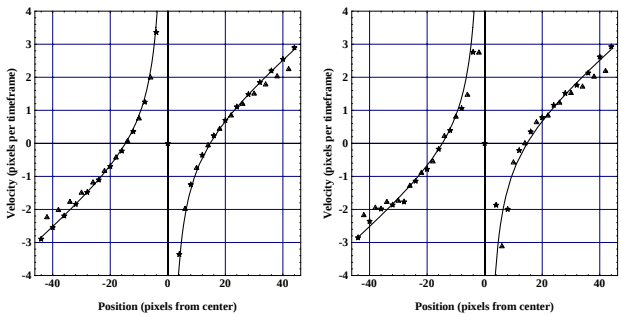

Figure 10: Cross-section of the velocity field obtained for the density stimulus of Formula 1; analytical (solid line), zeroth (triangles) and first order (stars). **Left:** noise free case. **Right:** 50% noise.

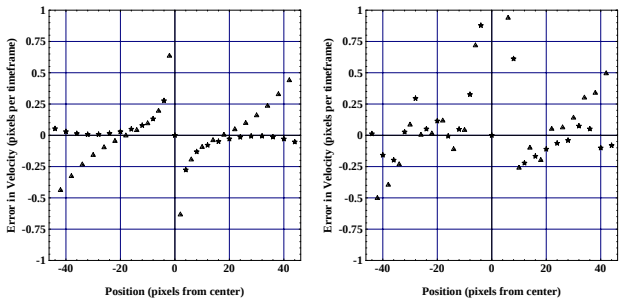

Figure 11: Error in cross-section of the velocity field obtained for the density stimulus of Formula 1; zeroth (triangles) and first order (stars). **Left:** noise free case. **Right:** 50% noise.

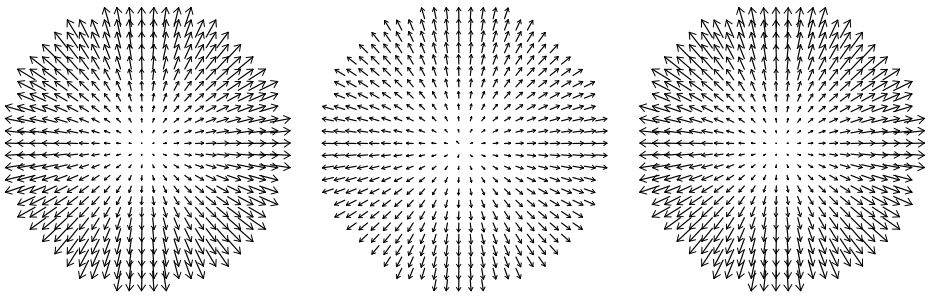

Figure 12: **Left**: Analytically determined velocity field for the scalar stimulus of Formula 9. **Middle & Right**: Velocity field obtained in the zeroth and first order approximation ($\sigma_s = 2; \sigma_t = 1$).

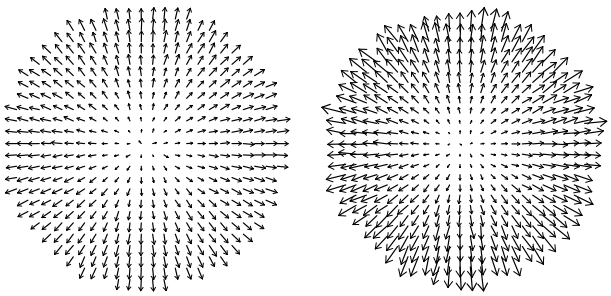

Figure 13: Velocity field obtained for the scalar stimulus of Formula 9. 50% Multiplicative Gaussian noise has been added to the original data. **Left**: zeroth order approximation. **Right**: first order approximation.

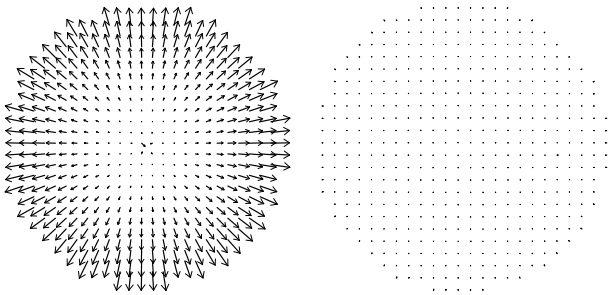

Figure 14: Error in velocity field obtained for the scalar stimulus of Formula 9. **Left**: zeroth order approximation. **Right**: first order approximation.

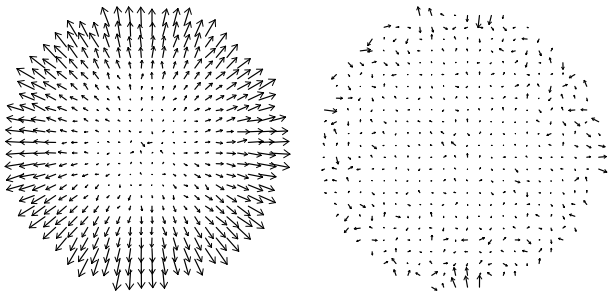

Figure 15: Error in velocity field obtained for the scalar stimulus of Formula 9, after the original image has been corrupted with 50% multiplicative Gaussian noise. **Left**: zeroth order approximation. **Right**: first order approximation.

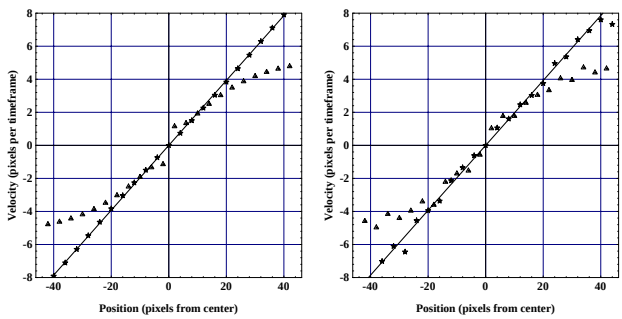

Figure 16: Cross-section of the velocity field obtained for the scalar stimulus of Formula 9; analytical (solid line), zeroth (triangles) and first order (stars). **Left:** noise free case. **Right:** 50% noise.

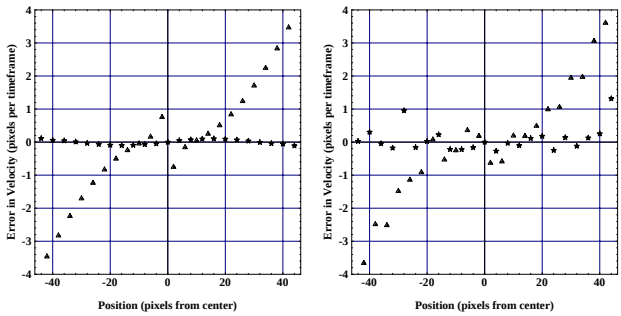

Figure 17: Error in cross-section of the velocity field obtained for the scalar stimulus of Formula 9; zeroth (triangles) and first order (stars). **Left:** noise free case. **Right:** 50% noise.

5 Comparison with Existing Techniques

In Section 5.1 we make a conceptual comparison with models that are most closely related to ours, notably by Otte and Nagel [68], and by Werkhoven and Koenderink [83]. In Section 5.2 we compare the performance of zeroth and first order optic flow schemes to those reported by Barron *et al.* [7], using an appropriate gauge condition.

5.1 Conceptual Comparison with Similar Methods

Upon first glance there is nothing remarkable about the linear systems in Section 4.3 (or Result 1 in general) when compared to similar ones proposed in the literature. It would lead too far to make a scrutinised comparison here. However, the similarity with a few of these, such as suggested by Otte and Nagel [68], or by Werkhoven and Koenderink [83], is particularly striking. It is therefore instructive to point out how our theory differs from theirs.

Among others, Otte and Nagel discuss two OFCE-based linear systems, which they refer to as the "Rigorous Condition" (RC) and the "Integrated Condition" (IC), respectively, and of which they have presented particular orders by way of example. The gauge invariant part of the RC (top four rows of [68, Equation 8]) is readily obtained from the first order system above by taking the limit $\sigma, \tau \downarrow 0$ for the *explicit* scale factors. This means that such factors as $L_x + L_{xtt}$ will effectively converge to L_x, *et cetera*. If the remaining derivatives are still computed at finite scales $\sigma, \tau > 0$, one obtains a conservation law that is assumed to hold *at these finite scales*. Recall that in our scheme, conservation applies to the hypothetical level $\sigma = \tau = 0$, whereas its consequences are investigated at finite measurement scales. The reader may verify that if the scalar OFCE is applied to the stimulus of Equation 9 after spatial blurring to scale $\sigma_0 > 0$, the flow will develop a singularity at the origin proportional to

$$\|\vec{v}(r)\| \propto \frac{s_0}{s} \frac{|\dot{s}|}{r} \qquad (r \downarrow 0)\,, \qquad \text{with } 2s_0 = \sigma_0^2\,,$$

which signifies the inadequacy of the RC. The rest of the RC (all other rows of [68, Equation 8]) has no counterpart in our theory, as we argued this to be part of a specific kind of gauge fixing.

More closely related to our work seems to be Otte and Nagel's IC ([68, Equations 10–11a]). Again, to be able to compare we must disregard their second order equations, which correspond to semantically inspired gauge conditions. In addition we must include temporal derivatives; altogether we must consider the analogous, underdetermined (gauge invariant) set of 4 equations in 8 unknowns, as in the $k = 1$ system above. In a way the IC is based on duality, which explains the similarity with our method. However, under the IC the (approximate) Lie derivative is integrated over a spatiotemporal cylinder (the temporal extension of the spatial sphere used in [68, Equation 9]), and not computed as a smoothly weighted integral. From a duality perspective, the IC-filter is an indicator function on a cylinder of spatial radius and temporal extent proportional to σ and τ, respectively, which prohibits transposition of derivatives. Recall that for the purpose of well-posed and operationally well-defined derivation,

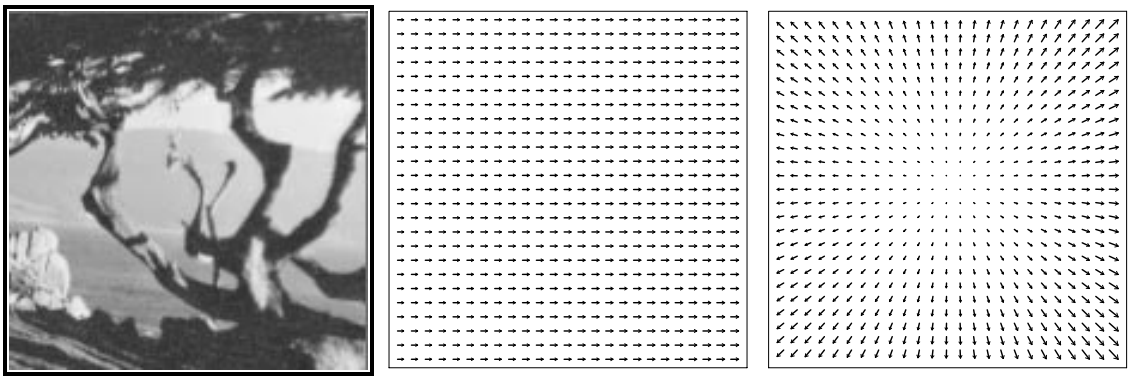

Figure 18: Planar textured surface and vector field for translational and diverging motion.

filter smoothness is an essential demand in our strategy. Taking into account the abovementioned modifications to the IC, one finds upon careful inspection that the coefficients resulting from the IC approach are indeed identical in form to the ones of the $k = 1$ system above.

In the context of visual motion detection Werkhoven and Koenderink do use the same duality principle as we do, based on the output of receptive fields from the Gaussian family [83]. In their article it is conjectured that

> "The aperture problem [...] is not an inherent problem in visual motion detection."

This is evidently true, since visual routines may well be gauge constrained. Curiously, although the authors do in fact propose an explicit and visually plausible mechanism for gauge fixing, *viz.* by monitoring the output of a *restricted* set of receptive field types, it does not seem to occur to them that they have *ipso facto* brought in an extrinsic source of information, as they adhere to the aforementioned deeply rooted *vox populi*:

> "Except for image irradiance patterns that vary only in one dimension, it [*i.e.* the aperture problem] does not arise if higher order spatial derivatives are considered."

Technically, however, Werkhoven and Koenderink's approach is probably the one that comes nearest to ours, as it is in fact a gauge constrained instance.

5.2 Performance Comparison with Similar Methods

Barron *et al.* have compared the performance of nine optic flow techniques reported in the literature [7]. In this section we evaluate zeroth and first order approximations for the two image sequences known as the *Translating Tree Sequence* (TTS) and the *Diverging Tree Sequence* (DTS), respectively [7, 16]: Figure 18. Notation: $\tilde{u}$ and $\tilde{v}$ denote horizontal and vertical components of the *ideal* velocity field, u, v are the zeroth order coefficients in the polynomial expansion, again without explicit reference to the value of approximation order M, and subscripts x and y denote corresponding derivatives or polynomial coefficients. In all that follows, scales are set to unity for the sake of notational simplicity: $\sigma = \tau = 1$; with a straightforward rescaling these parameters can be made explicit again.

Our point of departure will be Result 1, with $M = 0$ or $M = 1$. Recall that this result involves only order of approximation and formal geometric interpretation, *in casu* the scalar flow paradigm; apart from this, its form is independent of task or other external factors. Also recall that the price of this flexibility is an unresolved ambiguity in the form of local gauge invariance. To be able to connect to the conditions of the benchmark test—at the same time illustrating the intentional use—we therefore complement our intrinsic equations with knowledge of camera motion.

In the TTS case the camera moves in front of a textured plane along its x-axis, inducing horizontal velocities. In the DTS case the camera moves towards the plane along its line of sight with the *focus of*

expansion (FOE) in the centre of the image. Figure 18 shows the textured plane and the velocity for both cases[11].

For the TTS we use the additional constraint

$$\tilde{v} = 0\,, \tag{14}$$

which is reflected in the respective polynomial approximators as follows:

$$v = 0 \qquad (M = 0)\,, \tag{15}$$

$$v = v_x = v_y = v_t = 0 \qquad (M = 1)\,. \tag{16}$$

This suffices to solve for a unique, meaningful flow field.

In the case of the DTS the velocity will always be radial relative to the FOE. We can use this as an additional constraint:

$$x\tilde{v} - y\tilde{u} = 0\,, \tag{17}$$

with x and y relative to the FOE. In the zeroth order case ($M = 0$: $\tilde{u} = u, \tilde{v} = v$) this becomes

$$xv - yu = 0\,. \tag{18}$$

Combining this with Result 1 for $M = 0$, we can uniquely solve for the zeroth order velocity field.

As for first order, we may differentiate the gauge constraint 17 with respect to x, y and t:

$$\tilde{v} - y\tilde{u}_x + x\tilde{v}_x = 0\,, \tag{19}$$

$$-\tilde{u} - y\tilde{u}_y + x\tilde{v}_y = 0\,, \tag{20}$$

$$-y\tilde{u}_t + x\tilde{v}_t = 0\,, \tag{21}$$

which, upon substitution of a first order polynomial ($M = 1$: $\tilde{u} = u + xu_x + yu_y + tu_t, \tilde{v} = v + xv_x + yv_y + tv_t$), produces four constraints. Note that none of these are admissible in the zeroth order case. Altogether, the first order DTS system is given by

$$\begin{pmatrix} L_x & L_y & L_{xt} & L_{yt} & L_{xx} & L_{xy} & L_{xy} & L_{yy} \\ L_{xt} & L_{yt} & L_x + L_{xtt} & L_y + L_{ytt} & L_{xxt} & L_{xyt} & L_{xyt} & L_{yyt} \\ L_{xx} & L_{xy} & L_{xxt} & L_{xyt} & L_x + L_{xxx} & L_y + L_{xxy} & L_{xxy} & L_{xyy} \\ L_{xy} & L_{yy} & L_{xyt} & L_{yyt} & L_{xxy} & L_{xyy} & L_x + L_{xyy} & L_y + L_{yyy} \\ -y & x & -yt & xt & -xy & x^2 & -y^2 & xy \\ 0 & 0 & -y & x & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & t & -y & 2x & 0 & y \\ -1 & 0 & -t & 0 & -x & 0 & -2y & x \end{pmatrix} \begin{pmatrix} u \\ v \\ u_t \\ v_t \\ u_x \\ v_x \\ u_y \\ v_y \end{pmatrix} = \begin{pmatrix} -L_t \\ -L_{tt} \\ -L_{xt} \\ -L_{yt} \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{22}$$

Note that the first four rows express syntactical flow as given by Result 1 for $M = 1$, and retain their validity for the TTS sequence, whereas the last four rows reflect our assumption of camera movement. The FOE condition expressed by the last four rows in Equation 22 is actually a globally valid, *spatial* constraint. For each time frame we therefore have $t = 0$, because $(t; x, y)$ labels points relative to the *enduring* FOE at processing time; the entries $(5,3), (5,4), (7,4), (8,3)$ must therefore be set equal to zero. In the TTS case the semantical part of the matrix contains only 0's and 1's, the latter in entries $(5,2), (6,4), (7,6), (8,8)$, in accordance with Equation 16.

5.2.1 Scale Selection

The linear systems discussed previously form the basis of our experimental study, but note that one can plug in *arbitrary* measurement scales (we have chosen a few *ad hoc* values to play with). The question therefore arises of how to *select appropriate scales*. Intuitively one would want to choose a spatial scale corresponding to the size of features in a scene, and let velocities of objects determine an appropriate temporal scale (velocities are naturally expressed in terms of the ratio of spatial and temporal scale units). The possibility of scale selection is in fact a powerful feature of the theory.

We have implemented an *automatic scale selection* mechanism in which scales are played off against numerical stability. This procedure turns out to yield better performance than any fixed scale setting

[11] Data and flow fields have been obtained from the University of Western Ontario (csd.uwo.ca) by anonymous ftp.

we have tried. The method is the one proposed by Niessen *et al.* [64, 65, 66], and works as follows. Consider the linear system

$$\mathbf{A}\mathbf{v} = \mathbf{b}\,. \tag{23}$$

Generically this equation will be nonsingular, but if $\mathbf{A}$ depends on control parameters—*in casu* scale—a proper setting may be desirable from a numerical point of view. In particular, since both $\mathbf{A}$ as well as $\mathbf{b}$ are measurements with an intrinsic uncertainty, it is important to select values for which the system of Equation 23 is sufficiently well-conditioned.

A suitable measure readily available from standard matrix algebra is the so-called *condition number*, which relates the induced error in $\mathbf{v}$ to those in $\mathbf{A}$ and $\mathbf{b}$ (small values are preferred):

$$K(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|\,. \tag{24}$$

Here, $\|\,.\,\|$ denotes matrix norm. For instance, using a 2-norm we have

$$K(\mathbf{A}) = \frac{\lambda_+(\mathbf{A})}{\lambda_-(\mathbf{A})}\,, \tag{25}$$

in which $\lambda_+(\mathbf{A})$ is the largest and $\lambda_-(\mathbf{A})$ the smallest singular value of $\mathbf{A}$ (*i.e.*, eigenvalue of $\mathbf{A}^{\mathrm{T}}\mathbf{A}$). Apart from a small condition number we prefer a large signal to noise ratio. To this end we require the smallest singular value to be larger than a fiducial threshold λ [67]:

$$\lambda_-(\mathbf{A}) > \lambda\,. \tag{26}$$

Considering this we have chosen an approach which takes all singular values into account by selecting parameters that minimise the so-called *Frobenius norm* of $\mathbf{A}^{-1}$, defined as the sum of squares of all singular values:

$$\|\mathbf{A}^{-1}\|_{\mathrm{F}}^2 = \sum_\alpha \lambda_\alpha^2(\mathbf{A}^{-1}) = \sum_\alpha \frac{1}{\lambda_\alpha^2(\mathbf{A})}\,. \tag{27}$$

Note that if one of the singular values of $\mathbf{A}$ is small this norm will become large. Furthermore, if $\|\mathbf{A}^{-1}\|_{\mathrm{F}}^2$ is small, all singular values are large and the matrix is usually well-conditioned.

5.2.2 Results

We have carried out the computation at multiple spatial and temporal scales ($\sigma = 1.0, 1.414, 2.0, 2.828, 4.0, 5.656, 8.0$, and $\tau = 1.0, 2.0$, respectively). We have also computed a velocity field based on the scale selection criterion for these scales.

In order to be able to compare results we use the same angular error measure ε as proposed by Fleet and Jepson [16], and used by Barron *et al.* [7], which is based on the spatiotemporal orientation of the measured vector v_{e} relative to that of the correct vector v_{c} (recall that $v = (1; u, v)$):

$$\varepsilon = \arccos(\widehat{v}_{\mathrm{c}} \cdot \widehat{v}_{\mathrm{e}}) \qquad \text{with} \qquad \widehat{v} = \frac{v}{\sqrt{v \cdot v}}\,. \tag{28}$$

Although we have maintained confidence measures with velocity estimates, we have used a threshold on the Frobenius norm to discard uncertain vectors. In Table 1 the results are listed for various spatiotemporal scales. The following conclusions can be drawn:

- First order approximation does not necessarily perform better than zeroth order for *fixed* scales. Especially at fine scales the accuracy of higher order derivatives (third order is needed in first order approach) turns out to be problematic.
- Scale selection significantly improves results if all vectors are kept. If for a properly chosen spatial and temporal scale uncertain vectors are discarded, single scale methods may be as good as or even outperform the one based on automatic scale selection.
- Scale selection tends to remove outliers, which is apparent from the decreasing standard deviation.

Parameters OFCE	TTS mean	TTS st. dev.	TTS density	DTS mean	DTS st. dev.	DTS density
$M = 0, \sigma = 1, \tau = 2$	3.17	13.04	100%	3.50	10.29	100 %
$M = 1, \sigma = 1, \tau = 2$	3.89	15.11	100%	8.14	17.44	100 %
$M = 0, \sigma = 1, \tau = 2$	0.61	1.08	60%	0.98	1.27	60 %
$M = 1, \sigma = 1, \tau = 2$	0.70	1.35	60%	2.28	3.88	60 %
$M = 0, \sigma = 2, \tau = 2$	1.60	8.68	100%	3.88	10.65	100 %
$M = 1, \sigma = 2, \tau = 2$	1.50	8.82	100%	4.76	12.69	100 %
$M = 0, \sigma = 2, \tau = 2$	0.38	0.58	60%	1.29	1.53	60 %
$M = 1, \sigma = 2, \tau = 2$	0.29	0.41	60%	1.32	2.06	60 %
$M = 0, \sigma = 4, \tau = 2$	1.39	8.48	100%	7.39	14.39	100 %
$M = 1, \sigma = 4, \tau = 2$	0.71	5.73	100%	2.65	8.63	100 %
$M = 0, \sigma = 4, \tau = 2$	0.33	0.43	60%	2.66	2.68	60 %
$M = 1, \sigma = 4, \tau = 2$	0.16	0.18	60%	0.79	1.13	60 %
$M = 0, \sigma = 8, \tau = 2$	2.08	10.03	100%	11.98	19.65	100 %
$M = 1, \sigma = 8, \tau = 2$	1.43	8.53	100%	3.32	9.61	100 %
$M = 0, \sigma = 8, \tau = 2$	0.79	3.65	60%	4.35	3.99	60 %
$M = 1, \sigma = 8, \tau = 2$	0.42	1.04	60%	1.56	3.88	60 %
$M = 0$, multiple σ, τ	0.57	1.76	100%	4.99	7.93	100 %
$M = 1$, multiple σ, τ	0.49	1.92	100%	1.15	3.32	100 %
$M = 0$, multiple σ, τ	0.42	0.88	60%	2.68	2.66	60 %
$M = 1$, multiple σ, τ	0.28	0.65	60%	0.92	2.51	60 %

Table 1: Error statistics for fixed-scale and multi-scale techniques for the OFCE based on the error measure ε (28).

We have compared the performance of our algorithm with the results reported by Barron *et al.* [7] (for details the reader is referred to the literature). Table 2 concerns methods in which all vectors are maintained ("density = 100%"). Our OFCE method compares favourably with these approaches. This should not come as a complete surprise, since adequate *a priori* knowledge has been brought in. However, the results provide support for precisely the kind of approach we have adopted, based on a reasonable and fairly uncommitted syntactical definition of optic flow combined with a clear entry point for sophisticated semantics. It is also evident that scale selection significantly improves results. Finally, in Table 3 we compare our approach with the best performing algorithms that discard uncertain estimates. Our approach performs better than the algorithm of Fleet & Jepson, which is the best performing algorithm reported [7].

Implementation method	TTS mean	TTS st. dev.	DTS mean	DTS st. dev.
Horn & Schunck (original)	38.72	27.67	12.02	11.72
Modified Horn & Schunck	2.02	2.27	2.55	3.67
Uras *et al.* (unthresholded)	0.62	0.52	4.64	3.48
Nagel	2.44	3.06	2.94	3.23
Anandan	4.54	3.10	7.64	4.96
Singh (step 1, $n = 2, w = 2, (N = 4)$)	1.64	2.44	17.66	14.25
Singh (step 2, $n = 2, w = 2, (N = 4)$)	1.25	3.29	8.60	5.60
OFCE ($M = 0, \sigma = 2, \tau = 2$)	1.60	8.68	3.88	10.65
OFCE ($M = 1, \sigma = 4, \tau = 2$)	0.71	5.73	2.65	8.63
OFCE ($M = 1$, multiple σ, τ)	0.49	1.92	1.15	3.32

Table 2: OFCE compared with best performing techniques [7] with velocity estimates in the entire image domain. Again, scale selection involves the two temporal and seven spatial scale values previously mentioned.

Implementation method	TTS			DTS		
	mean	st. dev.	density	mean	st. dev.	density
Horn & Schunck (original, $\|\nabla L\| \geq 5.0$)	32.66	24.50	55.9%	8.93	7.79	54.8 %
Modified Horn & Schunck	1.89	2.40	53.2%	1.94	3.89	32.9 %
Lucas and Kanade ($\lambda_2 \geq 1.0$)	0.66	0.67	39.8%	1.94	2.06	48.2 %
Lucas and Kanade ($\lambda_2 \geq 5.0$)	0.56	0.58	13.1%	1.65	1.48	24.3 %
Uras et al. ($\det(H) \geq 1.0$)	0.46	0.35	41.8%	3.83	2.19	60.2 %
Nagel $\|\nabla L\|_2 \geq 5.0$	2.24	3.31	53.2%	3.21	3.43	53.5 %
Singh (step 1, $n = 2, w = 2, \lambda_1 \leq 5.0, (N = 4)$)	0.72	0.75	41.4 %	7.09	6.59	3.3 %
Heeger	4.53	2.41	57.8%	4.49	3.10	74.2 %
Fleet & Jepson ($\tau = 2.5$)	0.32	0.38	74.5%	0.99	0.78	61.0 %
Fleet & Jepson ($\tau = 2.0$)	0.23	0.19	49.7%	0.80	0.73	46.5 %
Fleet & Jepson ($\tau = 1.0$)	0.25	0.21	26.8%	0.73	0.46	28.2 %
OFCE ($M = 0, \sigma = 4, \tau = 2$)	0.33	0.43	60%	2.66	2.68	60 %
OFCE ($M = 1$, multiple σ, τ)	0.16	0.18	60%	0.79	1.13	60 %
OFCE ($M = 0$, multiple σ, τ)	0.29	0.33	40%	2.04	1.81	40 %
OFCE ($M = 1$, multiple σ, τ)	0.14	0.13	40%	0.43	0.40	40 %

Table 3: OFCE compared with best performing techniques reported by Barron *et al.* [7].

6 Summary and Discussion

In this article we have presented new conceptual and computational aspects of optic flow measurement. An important novelty is the incorporation of measurement duality in the traditional approach based on the Optic Flow Constraint Equation, which has led to a consistent representation of optic flow that accounts for the role of preprocessing filters. As a spin-off we have derived a scale-space for optic flow fully compatible with the conventional scale-space paradigm for grey-scale images. Moreover, a gauge theoretical approach has been adopted in order to make the distinction "structure" *versus* "meaning" manifest, *viz.* by modelling the former (data evidence) in terms of a system characterised by a local ambiguity (gauge invariance, basically the "aperture problem" in disguise), and the latter (models, hypotheses) in terms of a complementary set of gauge constraints. Accordingly we have argued in favour of a conceptual distinction between "syntactical" (data-defined, gauge invariant) and "semantical" (model-induced, gauge constrained) flow.

Syntax pertains to structure, which is itself of course a product of reasoning. In general one needs to decide on the geometric interpretation of the image data—the "proto-semantics" of image formation: scalar versus density, topological duality, flow-defining conservation principle, *et cetera*—in order to define the flow field's syntactical structure. In addition one needs to address the semantics of the problem by gauge fixing; examples are the normal flow condition, and even the usual assumption of conservation of topological detail. In practice one will often need more sophisticated gauges. We believe that the theory is particularly suited for modelling flow in medical imaging, but with some modifications it may well be of wider applicability, *e.g.* in the context of binocular stereo [63, 65].

As for the computational part, we have presented a robust, operational definition of syntactical optic flow in arbitrary dimensions and to arbitrary orders of approximation. The emphasis has been on the scalar case, but it has also been pointed out how to deal with densities.

The theory for the scalar case has been verified by means of an analytically tractable stimulus. Numerical simulation of optic flow extraction in $(2 + 1)$D shows qualitatively acceptable results in lowest order, and quite accurate quantitative results in first order approximation, even in the presence of substantial multiplicative noise. We have also compared the performance of zeroth and first order schemes on a benchmark sequence with similar ones proposed in the literature [7]. The first order case outperforms existing techniques if appropriate hypotheses are used to fix the gauge.

An aspect that has not been addressed here is *optimal order*; we have formulated the optic flow extraction method for arbitrary orders of approximation M, involving partial image derivatives of orders as high as $2M + 1$. Optimality of differential order is therefore a decisive factor; this depends on various details such as noise, discretisation characteristics, reconstruction quality, resolution of interest, *et cetera*.

Higher orders are quite likely to improve results even further as far as these quality criteria permit. This is to be expected because the formal expansion v_M converges to the actual optic flow defining generator v as $M \to \infty$, at least in theory. In practice it is useful to consider the Cauchy differences $v_{M+1} - v_M$, since these are actually computable up to some order M. From these one can infer the rate of successive refinement of the various optic flow parameters, and thus get an impression of the rate of convergence of the local optic flow field expansion. In turn this enables predictions concerning which orders are relevant and which can be savely ignored, depending on image structure, instead of committing oneself to a predetermined truncation order. Moreover, if the Cauchy differences do not decrease sufficiently fast, so that they become negligible before the optimal order prohibits further refinement, then we effectively have an optic flow *discontinuity*. This is of interest in segmentation tasks. Finally, knowledge of refinement is also important because all optic flow parameters, even the lowest order ones, are only reliably extracted if M-th order provides a sufficiently accurate local approximation. Only in that case will order of approximation be a good substitute for differential order, so that we can compare the coefficients of v_M^μ with the corresponding Taylor coefficients of v^μ. The principle of refinement is therefore important, because no *a priori* truncation order M will guarantee a globally acceptable approximation. We have made sure, however, that truncation will never introduce a "spurious gauge", so that we can bring in knowledge in an unbiased way.

A largely unsolved problem is how to handle the deep structure of scale-space, although promising approaches do exist in the form of coarse-to-fine principles, automatic scale selection mechanisms, and topological descriptions based on catastrophe theory. In our numerical study we have exploited the scale degree of freedom so as to obtain the most reliable high-resolution optic flow approximation by monitoring the condition number of the defining linear system as a function of spatiotemporal scale.

Last but not least, it should be noted that the *real* problem in practice is always one of semantics, *viz.* to establish and validate appropriate gauge conditions on the basis of *a priori* knowledge. Semantical optic flow quite likely requires a hermeneutic circle in which complementary cues are synchronised, since any initially hypothesised model may fail independent of data quality (*e.g.* flow-segmentation-registration). Results must be evaluated relative to task performance and multiple cue consistency, not to "ground truth".

A The Filters $\Phi^{\rho_1\cdots\rho_l}_{\mu_1\cdots\mu_k}$

Using the following lemma we can get rid of the derivative ∂_ρ in the integrand of Result 1:

Lemma 1 *Using parentheses to denote index symmetrisation, we have*

$$\partial_\mu \Phi^{\rho_1\cdots\rho_l}_{\mu_1\cdots\mu_k}(x) = -\Phi^{\rho_1\cdots\rho_l}_{\mu_1\cdots\mu_k\mu}(x) + \delta^{(\rho_l}_{\mu}\Phi^{\rho_1\cdots\rho_{l-1})}_{\mu_1\cdots\mu_k}(x)\,.$$

It is understood that $\Phi^{\rho_1\cdots\rho_{l-1}}_{\mu_1\cdots\mu_k} \equiv 0$ *if* $l = 0$.

The proof of this lemma is straightforward and will be omitted. Using this lemma we can rewrite Result 1:

Result 3 *See Result 1.*

$$\partial_{\mu_1\ldots\mu_k}\mathcal{L}_{v_M}F[\phi] = \sum_{l=0}^{M} v^{\mu}_{M;\rho_1\ldots\rho_l}\int dx\, f(x)\,\left[\Phi^{\rho_1\cdots\rho_l}_{\mu_1\ldots\mu_k\mu}(x) - \Phi^{\rho_1\cdots\rho_{l-1}}_{\mu_1\ldots\mu_k}(x)\,\delta^{\rho_l}_{\mu}\right]\,.$$

(Note that we do not need to make index symmetrisation for $\rho_1,\ldots,\rho_l$ explicit here; it is automatically achieved by virtue of symmetry of the coefficients $v^{\mu}_{M;\rho_1\ldots\rho_l}$.)

In order to express the overcomplete set of filters $\Phi^{\rho_1\cdots\rho_l}_{\mu_1\ldots\mu_k}$ in terms of Gaussian derivative filters $\phi_{\mu_1\ldots\mu_m}$, consider the following diagram.

$$\begin{array}{ccc}
\Phi^{\rho_1\cdots\rho_l}_{\mu_1\ldots\mu_k}(x) & \xrightarrow{\mathbf{F}} & \hat{\Phi}^{\rho_1\cdots\rho_l}_{\mu_1\ldots\mu_k}(\omega) \\
\star\downarrow & & \downarrow\star\star \\
\Phi^{\rho_1\cdots\rho_l}_{\mu_1\ldots\mu_k}(x) & \xleftarrow{\mathbf{F^{-1}}} & \hat{\Phi}^{\rho_1\cdots\rho_l}_{\mu_1\ldots\mu_k}(\omega)
\end{array}$$

Instead of simplifying directly in the spatial domain (the act of which is symbolised by $\star$), we take the equivalent Fourier detour ($\mathbf{F} \to \star\star \to \mathbf{F^{-1}}$), and simplify in Fourier space ($\star\star$). We can make the following formal identifications of operators (the l.h.s. in the spatial domain, the r.h.s. in the Fourier domain):

$$x^\rho \xrightarrow{\mathbf{F}} i\frac{\partial}{\partial\omega_\rho}\,, \qquad \frac{\partial}{\partial x^\rho} \xrightarrow{\mathbf{F}} i\omega_\rho\,. \tag{29}$$

We need one more definition.

Definition 6 (Hermite Polynomials) *The Hermite polynomial of order* k, H_k, *is defined by*

$$\frac{d^k}{dx^k}\,e^{-\frac{1}{2}x^2} = (-1)^k H_k(x)\,e^{-\frac{1}{2}x^2}\,.$$

This is appropriate for the 1-dimensional case. Let us define the n-dimensional analogue of the Hermite polynomials as follows.

Definition 7 (Hermite Polynomials in n Dimensions) *The* n*-dimensional Hermite polynomial of order* k, $\mathcal{H}_{i_1\ldots i_k}$, *is defined by*

$$\frac{\partial^k}{\partial x^{i_1}\ldots\partial x^{i_k}}\,e^{-\frac{1}{2}x^2} = (-1)^k\mathcal{H}_{i_1\ldots i_k}(x)\,e^{-\frac{1}{2}x^2}\,.$$

These n-dimensional Hermite polynomials are related to the standard ones in the following way.

Lemma 2 (Relation to Standard Definition) *The* n*-dimensional Hermite polynomials as defined in Definition 7 are tensor products of one-dimensional Hermite polynomials as defined in Definition 6. More precisely,*

$$\mathcal{H}_{i_1\ldots i_k}(x) = \prod_{j=1}^{n} H_{\alpha^{i_1\ldots i_k}_j}(x^j)\,,$$

in which $\alpha^{i_1\ldots i_k}_j$ *denotes the number of indices in* $i_1,\ldots,i_k$ *equal to* j.

Clearly we have $\sum_{j=1}^{n} \alpha_j^{i_1 \ldots i_k} = k$, since this simply sums up all indices. The separability property of Lemma 2 follows straightforwardly from Definition 6, when applied to a multidimensional Gaussian.

Having established all basic ingredients and notational matters, we can now relate the overcomplete family of filters $\Phi_{\mu_1 \ldots \mu_k}^{\rho_1 \ldots \rho_l}$ to the Gaussian family. This is easy, since all we need to do is to use Leibniz product rule for differentiation in

$$\hat{\Phi}_{\mu_1 \ldots \mu_k}^{\rho_1 \ldots \rho_l}(\omega) = \frac{(-1)^k}{l!}\, i\frac{\partial}{\partial \omega_{\rho_1}} \ldots i\frac{\partial}{\partial \omega_{\rho_l}} \left(i\omega_{\mu_1} \ldots i\omega_{\mu_k} \hat{\phi}(\omega) \right) , \tag{30}$$

(see Formula 29 and the definition of the filters in Result 1). Then, each time we have to take a derivative of $\hat{\phi}(\omega)$, we use the explicit property of the Gaussian stated in Definition 7. In this way we arrive at

Result 4 (The Filters $\Phi_{\mu_1 \ldots \mu_k}^{\rho_1 \ldots \rho_l}$ and the Gaussian Family) *Let $\mathcal{S}$ denote the index symmetrisation operator (applying both to upper as well as lower indices), then we have*

$$\hat{\Phi}_{\mu_1 \ldots \mu_k}^{\rho_1 \ldots \rho_l}(\omega) = \frac{(-1)^k}{l!}\, \mathcal{S} \left\{ \sum_{m=0}^{\min(k,l)} (-1)^m \binom{l}{m} \frac{k!}{(k-m)!} \delta_{\mu_1}^{\rho_1} \ldots \delta_{\mu_m}^{\rho_m}\, i\omega_{\mu_{m+1}} \ldots i\omega_{\mu_k} (-i)^{l-m} \mathcal{H}^{\rho_{m+1} \ldots \rho_l}(\omega)\, \hat{\phi}(\omega) \right\} .$$

Fourier inversion yields

$$\Phi_{\mu_1 \ldots \mu_k}^{\rho_1 \ldots \rho_l}(x) = \frac{(-1)^k}{l!}\, \mathcal{S} \left\{ \sum_{m=0}^{\min(k,l)} (-1)^m \binom{l}{m} \frac{k!}{(k-m)!} \delta_{\mu_1}^{\rho_1} \ldots \delta_{\mu_m}^{\rho_m}\, \partial_{\mu_{m+1}} \ldots \partial_{\mu_k} i^{l-m} \mathcal{H}^{\rho_{m+1} \ldots \rho_l}(i\nabla)\, \phi(x) \right\} .$$

Note that this expression is real in the spatial domain, since $i^p \mathcal{H}^{\rho_1 \ldots \rho_p}(i\nabla)$ is a real differential operator for any $p \in \mathbb{Z}_0^+$. To see this, look at the explicit form of a Hermite polynomial:

$$H_k(x) = \sum_{m=0}^{[k/2]} (-1)^m \binom{k}{2m} (2m-1)!!\, x^{k-2m} , \tag{31}$$

in which $[x]$ denotes the integer part of $x \in \mathbb{R}$, *i.e.* the largest integer less than or equal to x, and in which the double factorial $(2m-1)!!$ indicates the product $1 \times 3 \times \ldots \times (2m-1)$. Consequently,

$$i^k H_k(i\frac{d}{dx}) = (-1)^k \sum_{m=0}^{[k/2]} \binom{k}{2m} (2m-1)!! \frac{d^{k-2m}}{dx^{k-2m}} , \tag{32}$$

very real indeed. The general n-dimensional case follows from this observation. Note also that the r.h.s. of Result 4 is a linear combination of Gaussian derivatives of the type $\phi_{\mu_1 \ldots \mu_p}$, with $p = 0, \ldots, k+l$. Thus we have indeed proven overcompleteness of the (apparently $(k+l)$-th order) filters $\Phi_{\mu_1 \ldots \mu_k}^{\rho_1 \ldots \rho_l}$ by explicitly rewriting them in terms of Gaussian derivatives.

Acknowledgement

The theoretical part of this work has been carried out partly at INRIA Sophia-Antipolis, France, and partly at INESC Aveiro, Portugal, as part of the ERCIM Fellowship Programme, financed by the Commission of the European Communities. Olivier Faugeras is gratefully acknowledged for a clarifying discussion on the subject. Software development has been done by the Computer Vision Research Group, Utrecht, The Netherlands. Experimental verification as described in the examples has been carried out at Yale University, New Haven, USA, and at the Computer Vision Research Group, Utrecht, The Netherlands. The final version of the manuscript has been prepared at DIKU Copenhagen, Denmark, and at the Department of Mathematics and Computer Science of the University of Utrecht, The Netherlands. A summary has been presented at the Computer Vision and Applied Geometry workshop in Nordfjordeid, Norway; the anonymous referees have provided me with detailed suggestions that have also been taken into consideration in this article (notably Figure 2).

References

[1] A. A. Amini. A scalar function formulation for optical flow. In Eklundh [13], pages 125–131.

[2] J. Arnspang. Notes on local determination of smooth optic flow and the translational property of first order optic flow. Technical Report 88/1, Department of Computer Science, University of Copenhagen, 1988.

[3] J. Arnspang. Optic acceleration. Local determination of absolute depth and velocity, time to contact and geometry of an accelerating surface. Technical Report 88/2, Department of Computer Science, University of Copenhagen, 1988.

[4] J. Arnspang. Motion constraint equations in vision calculus, April 1991. Dissertation for the Danish Dr. scient. degree, University of Utrecht, Utrecht, The Netherlands.

[5] J. Arnspang. Motion constraint equations based on constant image irradiance. *Image and Vision Computing*, 11(9):577–587, November 1993.

[6] J. Babaud, A. P. Witkin, M. Baudin, and R. O. Duda. Uniqueness of the gaussian kernel for scale-space filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):26–33, 1986.

[7] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, 1994.

[8] N. Bourbaki. *Éléments de Mathématique, Livre V: Espaces Vectoriels Topologiques*. Hermann, Paris, 1964.

[9] Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick. *Analysis, Manifolds, and Physics. Part I: Basics*. Elsevier Science Publishers B.V. (North-Holland), Amsterdam, 1991.

[10] P. Delogne, editor. *Proceedings of the International Conference on Image Processing 1996*, Lausanne, Switzerland, September 16–19 1996. IEEE.

[11] V. Devlaminck and J. Dubus. Estimation of compressible or incompressible deformable motions for density images. In Delogne [10], pages 125–128.

[12] J. D'Haeyer. Determining motion of image curves from local pattern changes. *Computer Vision, Graphics, and Image Processing*, 34:166–188, 1986.

[13] J.-O. Eklundh, editor. *Proceedings of the 3rd European Conference on Computer Vision*, number 800–801 in Lecture Notes in Computer Science, Berlin, 1994. Springer-Verlag.

[14] O. D. Faugeras and T. Papadopoulo. A theory of the motion fields of curves. *International Journal of Computer Vision*, 10(2):125–156, 1993.

[15] J. M. Fitzpatrick. The existence of geometrical density-image transformations corresponding to object motion. *CVGIP: Image Understanding*, 44:155–174, 1988.

[16] D. J. Fleet and A. D. Jepson. Computation of component image velocity from local phase information. *International Journal of Computer Vision*, 5(1):77–104, 1990.

[17] L. M. J. Florack. Grey-scale images. Technical Report ERCIM-09/95-R039, INESC Aveiro, Portugal, September 1995. URL: http://www-ercim.inria.fr/publication/technical_reports.

[18] L. M. J. Florack. Data, models, and images. In Delogne [10], pages 469–472.

[19] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever. Scale and the differential structure of images. *Image and Vision Computing*, 10(6):376–388, July/August 1992.

[20] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever. Cartesian differential invariants in scale-space. *Journal of Mathematical Imaging and Vision*, 3(4):327–348, November 1993.

[21] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever. Images: Regular tempered distributions. In Y.-L. O, A. Toet, H. J. A. M. Heijmans, D. H. Foster, and P. Meer, editors, *Proceedings of the NATO Advanced Research Workshop Shape in Picture - Mathematical Description of Shape in Greylevel Images*, volume 126 of *NATO ASI Series F: Computer and Systems Sciences*, pages 651–660, Berlin, 1994. Springer-Verlag.

[22] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever. Linear scale-space. *Journal of Mathematical Imaging and Vision*, 4(4):325–351, 1994.

[23] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever. The Gaussian scale-space paradigm and the multiscale local jet. *International Journal of Computer Vision*, 18(1):61–75, April 1996. Erratum: Fig. 3 is upside down.

[24] L. M. J. Florack and M. Nielsen. The intrinsic structure of the optic flow field. Technical Report ERCIM-07/94-R033 or INRIA-RR-2350, INRIA Sophia-Antipolis, France, July 1994. URL: http://www-ercim.inria.fr/publication/technical_reports.

[25] L. M. J. Florack, W. J. Niessen, and M. Nielsen. The intrinsic structure of optic flow incorporating measurement duality. Accepted for publication in the International Journal of Computer Vision, 1996.

[26] J. J. Gibson. *The Perception of the Visual World.* Houghton-Mifflin, Boston, 1950.

[27] R. Gilmore. *Catastrophe Theory for Scientists and Engineers.* Dover Publications, Inc., New York, 1993. Originally published by John Wiley & Sons, New York, 1981.

[28] E. C. Hildreth. *The Measurement of Visual Motion.* MIT Press, Cambridge, 1983.

[29] E. C. Hildreth. Computations underlying the measurement of visual motion. *Artificial Intelligence*, 23:309–354, 1984.

[30] B. K. P. Horn. *Robot Vision.* MIT Press, Cambridge, 1986.

[31] B. K. P. Horn and B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[32] B. K. P. Horn and B. G. Schunck. "determining optical flow": a retrospective. *Artificial Intelligence*, 59:81–87, 1993.

[33] P. Johansen. On the classification of toppoints in scale space. *Journal of Mathematical Imaging and Vision*, 4(1):57–67, 1994.

[34] J. J. Koenderink. The concept of local sign. In A. J. van Doorn, W. A. van de Grind, and J. J. Koenderink, editors, *Limits in Perception*, pages 495–547. VNU Science Press, Utrecht, 1984.

[35] J. J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.

[36] J. J. Koenderink. Optic flow. *Vision Research*, 26(1):161–180, 1986.

[37] J. J. Koenderink. Design for a sensorium. In W. von Seelen, G. Shaw, and U. M. Leinhos, editors, *Organization of Neural Networks - Structures and Models*, pages 185–207. VCH Verlagsgesellschaft mbH, Weinheim, Germany, 1988.

[38] J. J. Koenderink. Design principles for a front-end visual system. In R. Eckmiller and Ch. v. d. Malsburg, editors, *Neural Computers*, volume 41 of *NATO ASI Series F: Computer and Systems Sciences*, pages 111–118. Springer-Verlag, Berlin, 1988.

[39] J. J. Koenderink. A hitherto unnoticed singularity of scale-space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(11):1222–1224, November 1989.

[40] J. J. Koenderink. The brain a geometry engine. *Psychological Research*, 52:122–127, 1990.

[41] J. J. Koenderink. Mapping formal structures on networks. In T. Kohonen, K. Mäkisara, O. Simula, and J. Kangas, editors, *Artificial Neural Networks*, pages 93–98. Elsevier Science Publishers B.V. (North-Holland), 1991.

[42] J. J. Koenderink and A. J. van Doorn. Logical stratification of organic intelligence. In R. Trappl, editor, *Cybernetics and Systems*, pages 871–878. D. Reidel Publishing Company, 1986.

[43] J. J. Koenderink and A. J. van Doorn. Representation of local geometry in the visual system. *Biological Cybernetics*, 55:367–375, 1987.

[44] J. J. Koenderink and A. J. van Doorn. The basic geometry of a vision system. In R. Trappl, editor, *Cybernetics and Systems*, pages 481–485. Kluwer Academic Publishers, 1988.

[45] J. J. Koenderink and A. J. van Doorn. Receptive field families. *Biological Cybernetics*, 63:291–298, 1990.

[46] J. J. Koenderink, A. Kappers, and A. J. van Doorn. Local operations: The embodiment of geometry. In G. A. Orban and H.-H. Nagel, editors, *Artificial and Biological Vision Systems*, Basic Research Series, pages 1–23. Springer Verlag, Berlin, 1992.

[47] J. J. Koenderink and A. J. van Doorn. Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica Acta*, 22(9):773–791, 1975.

[48] J. J. Koenderink and A. J. van Doorn. Operational significance of receptive field assemblies. *Biological Cybernetics*, 58:163–171, 1988.

[49] J. J. Koenderink and A. J. van Doorn. Receptive field assembly pattern specificity. *Journal of Visual Communication and Image Representation*, 3(1):1–12, 1992.

[50] J. J. Koenderink and A. J. van Doorn. Second-order optic flow. *Journal of the Optical Society of America-A*, 9(4):530–538, April 1992.

[51] D. N. Lee. A theory of visual control of braking based on information about time-to-collision. *Perception*, 5:437–457, 1976.

[52] D. N. Lee. The optic flow field: the foundation of vision. *Phil. Trans. R. Soc. Lond.* B, 290:169–179, 1980.

[53] T. Lindeberg. Scale-space for discrete signals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(3):234–245, 1990.

[54] T. Lindeberg. Scale-space behaviour of local extrema and blobs. *Journal of Mathematical Imaging and Vision*, 1(1):65–99, March 1992.

[55] T. Lindeberg. Detecting salient blob-like image structures and their scale with a scale-space primal sketch: A method for focus-of-attention. *International Journal of Computer Vision*, 11(3):283–318, 1993.

[56] T. Lindeberg. Effective scale: A natural unit for measuring scale-space lifetime. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(10):1068–1074, October 1993.

[57] T. Lindeberg. *Scale-Space Theory in Computer Vision*. The Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers, 1994.

[58] T. Lindeberg and J.-O. Eklundh. On the computation of a scale-space primal sketch. *Journal of Visual Communication and Image Representation*, 2(1):55–78, 1990.

[59] D. C. Marr. *Vision*. Freeman, San Francisco, CA, 1982.

[60] H.-H. Nagel. On the estimation of optical flow: Relations between different approaches and some new results. *Artificial Intelligence*, 33:299–324, 1987.

[61] H.-H. Nagel. Direct estimation of optical flow and its derivatives. In G. A. Orban and H.-H. Nagel, editors, *Artificial and Biological Vision Systems*, Basic Research Series, pages 193–224. Springer Verlag, Berlin, 1992.

[62] H.-H. Nagel and W. Enkelmann. An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):565–593, 1986.

[63] M. Nielsen, R. Maas, W. J. Niessen, L. M. J. Florack, and B. M. ter Haar Romeny. Binocular stereo from grey-scale images. Technical Report 96/17, Department of Computer Science, University of Copenhagen, 1996.

[64] W. J. Niessen, J. S. Duncan, M. Nielsen, L. M. J. Florack, B. M. ter Haar Romeny, and M. A. Viergever. A multi-scale approach to image sequence analysis. *Computer Vision and Image Understanding*, 65(2), 1997. To appear.

[65] W. J. Niessen and R. Maas. Optic flow and stereo. In J. Sporring, M. Nielsen, L. M. J. Florack, and P. Johansen, editors, *Gaussian Scale-Space*, Computational Imaging and Vision Series, chapter 4. Kluwer Academic Publishers, 1996. To appear.

[66] W. J. Niessen, M. Nielsen, L. M. J. Florack, R. Maas, B. M. ter Haar Romeny, and M. A. Viergever. Multiscale optic flow using physical constraints. In P. Johansen, editor, *Proceedings of the Copenhagen Workshop on Gaussian Scale-Space Theory*, pages 72–79, May 10–13 1996. DIKU Tech. Rep. Nr. 96/19.

[67] M. Okutomi and T. Kanade. A locally adaptive window for signal matching. *International Journal of Computer Vision*, 7(2):143–162, 1992.

[68] M. Otte and H.-H. Nagel. Optical flow estimation: Advances and comparisons. In Eklundh [13], pages 51–60.

[69] T. Poston and I. Steward. *Catastrophe Theory and its Applications*. Pitman, London, 1978.

[70] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C; the Art of Scientific Computing*. Cambridge University Press, Cambridge, 1988.

[71] B. G. Schunck. The motion constraint equation for optical flow. In *Proceedings of the 7th International Conference on Pattern Recognition*, pages 20–22, Montreal, Canada, 1984.

[72] L. Schwartz. *Théorie des Distributions*. Hermann, Paris, second edition, 1966.

[73] M. Spivak. *Calculus on Manifolds*. W. A. Benjamin, New York, 1965.

[74] M. Spivak. *Differential Geometry*, volume 1–5. Publish or Perish, Berkeley, 1975.

[75] M. Tistarelli. Multiple constraints for optical flow. In Eklundh [13], pages 61–70.

[76] M. Tistarelli. Multiple constraints to compute optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(12):1243–1250, December 1996.

[77] O. Tretiak and L. Pastor. Velocity estimation from image sequences with second order differential operators. In *Proceedings of the 7th International Conference on Pattern Recognition*, pages 16–19, Montreal, Canada, 1984.

[78] F. Trêves. *Topological Vector Spaces, Distributions and Kernels.* Academic Press, 1969.

[79] S. Uras, F. Girosi, A. Verri, and V. Torre. A computational approach to motion perception. *Biological Cybernetics*, 60:79–87, 1988.

[80] A. Verri, F. Girosi, and V. Torre. Differential techniques for optical flow. *Journal of the Optical Society of America-A*, 7(5):912–922, May 1990.

[81] A. Verri and T. Poggio. Motion field and optical flow: Qualitative properties. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(5):490–498, 1989.

[82] J. Weber and J. Malik. Robust computation of optical flow in a multi-scale differential framework. *International Journal of Computer Vision*, 14(1):67–81, 1995.

[83] P. Werkhoven and J. J. Koenderink. Extraction of motion parallax structure in the visual system I. *Biological Cybernetics*, 63:185–191, 1990.

[84] A. P. Witkin. Scale-space filtering. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 1019–1022, Karlsruhe, Germany, 1983.