

A Theory of Multiple Orientation Estimation

Matthias Mühlich and Til Aach

RWTH Aachen University, 52056 Aachen, Germany
{matthias.muehlich, til.aach}@lfb.rwth-aachen.de

Abstract. Estimation of local orientations in multivariate signals (including optical flow estimation as special case of orientation in space-time-volumes) is an important problem in image processing and computer vision. Modelling a signal using only a *single* orientation is often too restrictive, since occlusions and transparency happen frequently, thus necessitating the modelling and analysis of *multiple orientations*.

In this paper, we therefore develop a unifying mathematical model for multiple orientations: beyond describing an arbitrary number of orientations in multivariate vector-valued image data such as color image sequences, it allows the unified treatment of *transparently* and *occludingly* superimposed oriented structures. Based on this model, we derive novel estimation schemes for an arbitrary number of superimposed orientations in bivariate images as well as for double orientations in signals of arbitrary signal dimensionality. The estimated orientations themselves, but also features like the number of local orientations or the angles between multiple orientations (which are invariant under rotation) can be used for various inspection, tracking and segmentation problems. We evaluate the performance of our framework on both synthetic and real data.

1 Introduction: (Single) Orientation Estimation

Local orientations are an important low level feature for analyzing and understanding multivariate data. The basis for the concept of orientations is the important observation that signal gradients usually vary much slower than the signal itself. However, fast variations of gradients *do* appear in signals, for instance at corners in image data. In this paper, we want to promote the perspective that limitations of the orientation concept can be overcome by considering multiple signal orientations, thus making the concept of orientations even more fundamental for signal analysis. But first, let us review single orientation estimation.

1.1 The Structure Tensor

Let $\mathbf{x} \in \mathbb{R}^N$ be a vector in N -dimensional space. Then a function of \mathbf{x} , e.g. $s(\mathbf{x})$, defines a multivariate¹ signal. Such a signal is called locally *oriented* in some region Ω if it is constant along parallel lines, i.e.

¹ The signal s is scalar-valued and therefore *one-dimensional*. But it depends on a multidimensional vector and is therefore correctly called a *multivariate* (here: N -variate) signal. *Multidimensional* signals also exist in the orientation estimation context, for instance color images (which are three-dimensional and bivariate).

$$s(\mathbf{x} + \lambda \mathbf{u}) = s(\mathbf{x}) \quad \text{for all } \lambda \in \mathbb{R} \text{ and } \mathbf{x}, \mathbf{x} + \lambda \mathbf{u} \in \Omega \quad (1)$$

with some unit vector \mathbf{u} denoting the *orientation direction*.

A given signal is locally constant with respect to some unit vector \mathbf{u} , if its directional derivative $\frac{\partial s}{\partial \mathbf{u}} = \langle \mathbf{g}, \mathbf{u} \rangle$, i.e. the scalar product between signal gradient \mathbf{g} and \mathbf{u} , is zero for all gradients computed in some local neighborhood Ω . The gradients span a subspace in which the signal is *not* oriented and consequently, the orthogonal complement of this subspace is the sought orientation (which is uniquely determined if we can find $N - 1$ linearly independent gradients).

The introduction given so far is one out of several possible approaches leading to the so-called *structure-tensor* approach for orientation estimation which can be found in pioneering work of Förstner [1], Bigün et al [2] and others. For bivariate image data ($N = 2$; generalization to arbitrary N is straightforward), we first compute the discrete derivative of the signal with respect to x and y using convolution with filters f_x and f_y : $s_x = f_x * s$ and $s_y = f_y * s$. With the image gradient $\mathbf{g} = (s_x, s_y)^T$, we now define the (standard) structure tensor $\mathbf{S}^{(1)}$ as local integration over the outer product of the gradient:

$$\mathbf{S}^{(1)} = f' * (\mathbf{g}\mathbf{g}^T) \quad (2)$$

where f' is some averaging filter. (Widely used choices for f_x , f_y , and f' are (directional derivative of) Gaussian filters.) If (1) is only valid for a single orientation in general N -variate signals, then the structure tensor $\mathbf{S}^{(1)}$ has one zero eigenvalue and the corresponding eigenvector is the sought orientation. For noisy data or model violations, the eigenvector corresponding to the smallest eigenvalue defines the orientation in which the signal is “most constant”.

1.2 Related Approaches

The structure tensor $\mathbf{S}^{(1)}$ is not the only possibility to analyse single-oriented structures. As pointed out in [3], higher order directional derivatives also vanish in the orientation direction:

$$\frac{\partial s}{\partial \mathbf{u}} = 0 \quad \text{and} \quad \frac{\partial^2 s}{\partial \mathbf{u}^2} = 0 \quad \text{and} \quad \frac{\partial^3 s}{\partial \mathbf{u}^3} = 0 \quad \text{and} \quad \dots \quad (3)$$

which allows to design a wide class of approaches based on combinations of different order derivatives. This freedom can be used for filter design. For instance, the book of Granlund and Knutsson [4] gives a slightly different definition of the *orientation* concept: the *invariance requirement* states that an entity which characterizes orientation must not depend on the signal variations orthogonal to the sought orientation. This defines a much stronger concept of orientation than the one defined in (1). For instance in bivariate signals (i.e. images), it forces us to design an orientation estimator such that it makes no distinction between two especially important types of oriented 1D structures: “lines” (variation on two sides; also called “ridges”) and “edges” (variation only on one side). A detector that reacts uniformly on these two types of structures is called *phase invariant* and can be realized with *quadrature filters* [4] and a bandpass prefilter.

If, on the contrary, we are interested in detecting line structures only (and not edges), the second order directional derivative defines a popular filter; this can be traced back to [5]. A recent PAMI paper [6] also discusses line-specific and edge-specific orientation estimation in the context of steerable filters.

Two generalizations of the structure tensor to phase-invariant feature detectors are the *2D energy tensor* defined by Felsberg and Granlund in [7] and the *boundary tensor* proposed by Köthe in [8]. The connections between energy tensor and boundary tensor are analyzed recently by both authors in a joint paper [9]. All these approaches are based on higher (up to fourth) order derivatives.

Summarizing this section, we emphasize that the standard structure tensor approach can be extended with combinations of higher order derivatives in order to obtain advantageous properties. Odd order filters can be optimized for edge detection, even order filters for lines, and mixed order filters for phase invariant behaviour.

2 Modelling and Estimation of Multiple Orientations

Higher order derivatives in the context of orientation estimation also appear in a different line of research: the analysis of *multiple* orientations. In spite of characterizing many important low level image features like lines or edges, the underlying single orientation signal model is much too restrictive for many real signals. For instance, the presence of two oriented textures in a region Ω calls for an extended mathematical model. This observation led to the study of *double*-oriented signals (we will denote the multiplicity of orientations by M , so double orientations estimation means $M = 2$).

For image sequences, double orientation estimation means the study of two independent optical flows; this is the area where double orientation estimation appeared first in the beginning of the 90s in pioneering work of Shizawa/Mase [10, 11] (additive superposition model, grey value image sequences), followed by Shizawa/Iso [12] (additive superposition, grey value images, connection to steerable filters). More recent results can be found in [13, 14] (additive model, images; multispectral signals in the second reference; first theoretical steps towards higher multiplicity of signals beyond double orientations) and [15] (occluding model, multispectral images). However, present day algorithms are still limited to the estimation of double orientations ($M = 2$) in image or volume data ($N \leq 3$). Summarizing the previous work, we emphasize that *model* (additive or occluding), *signal dimensionality* (grey value or multispectral), *N -variate signals* (bivariate images, volumes, $N > 3$), and *multiplicity M* (double orientations or $M > 2$) define four ‘orthogonal’ directions for extensions of the early multiple orientations estimation work. First steps in all directions have been made, but the unifying theory is still missing. Most importantly, no experiments for triple or more orientations ($M \geq 3$) can be found anywhere in literature.

2.1 Transparent and Occluding Orientations

Two different ways of combining two or more oriented signals s_i (with $i = 1, \dots, M$) to form a new signal s can be found in literature. The *occluding orientations model* (OOM) and the *transparent orientations model* (TOM):

Multiple orientation models:

$$\text{(OOM:)} \quad s(\mathbf{x}) = s_i(\mathbf{x}) \quad \forall \mathbf{x} \in \Omega_i \quad \text{(TOM:)} \quad s(\mathbf{x}) = \sum_{i=1}^M \alpha_i s_i(\mathbf{x}). \quad (4)$$

The first model states that we take the first oriented signal if the point \mathbf{x} is in some region Ω_1 and so on (obviously, all regions must be distinct and add up to the whole analysis region Ω). For instance, this model is applicable with $M = 2$ if the region Ω_1 corresponds to some object which occludes another object (Ω_2 : background), provided that both objects can be modelled reasonably well as single-oriented structures.

The second model, TOM, assumes that all basic signals are present in the whole signal and we observe a superposition of them, weighted with some constants α_i .² Computing directional derivatives, we obtain the constraints:

Multiple orientation constraints, derivative forms:

$$\text{(OOM:)} \quad \prod_{i=1}^M \frac{\partial s}{\partial \mathbf{u}_i} = 0 \quad \text{(TOM:)} \quad \frac{\partial^M s}{\partial \mathbf{u}_1 \cdots \partial \mathbf{u}_M} = 0. \quad (5)$$

The directional derivative is defined as

$$\langle \mathbf{g}, \mathbf{u}_i \rangle = \frac{\partial s}{\partial \mathbf{u}_i} = \langle \nabla s, \mathbf{u}_i \rangle = \sum_{j=1}^N \frac{\partial s}{\partial x_j} (\mathbf{u}_i)_j \quad (6)$$

and inserting in the left hand side of (5) yields

$$\prod_{i=1}^M \left(\sum_{j=1}^N \frac{\partial s}{\partial x_j} (\mathbf{u}_i)_j \right) = \sum_{k_1, \dots, k_M=1}^N (\mathcal{O})_{k_1 \dots k_M} (\mathcal{U})_{k_1 \dots k_M} = \langle \mathcal{O}, \mathcal{U} \rangle = 0 \quad (7)$$

where we have rewritten a product of M factors (which are sums consisting of N summands each) as a large sum of N^M summands and then as scalar product of two tensors (i.e. sum over all element-by-element products). The whole dependency on the sought orientations is encapsulated in the tensor

$$\mathcal{U} = \mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_N \quad (8)$$

² It is also possible to define the weights α_i as functions of \mathbf{x} (instead of constants). Then, the OOM is a subset of the (generalized) TOM. Therefore, the TOM-approach can also be used for estimation under the OOM, though at the expense of needing higher order derivatives.

where “ \otimes ” denotes the tensor product operator. We now can state the multiple orientation constraints for both models. We first obtain the

Multiple occluding orientations constraint:

$$\langle \mathcal{O}, \mathcal{U} \rangle = 0 \quad \text{with} \quad (\mathcal{O})_{k_1 \dots k_M} = \prod_{i=1}^M \frac{\partial s}{\partial x_{k_i}} . \quad (9)$$

Each point in Ω yields one *data tensor* \mathcal{O} and from all these tensors, we have to estimate the sought orientation tensor \mathcal{U} which is orthogonal (i.e. has scalar product zero) to the given data tensors. Both \mathcal{O} and \mathcal{U} are $N \times \dots \times N$ tensors (all indices k_i with $i = 1, \dots, M$ run from 1 to N). For instance, in image sequences (trivariate data, $N = 3$), \mathcal{O} contains all possible products of M first order derivatives w.r.t. x , y and t coordinates. Analogously, we find the

Multiple transparent orientations constraint:

$$\langle \mathcal{T}, \mathcal{U} \rangle = 0 \quad \text{with} \quad (\mathcal{T})_{k_1 \dots k_M} = \frac{\partial^M s}{\partial x_{k_1} \dots \partial x_{k_M}} . \quad (10)$$

At the end of this subsection, we want to stress the structural similarity of both models: the tensors constructed from signal derivatives are different (product of first derivatives in \mathcal{O} versus higher order derivatives in \mathcal{T}), but once we have constructed the data tensor, the computation of the sought orientations (the estimation and decomposition of the orientation tensor \mathcal{U}) is *exactly the same*.

2.2 Symmetry Properties of the Data Tensors

The commutativity in the definitions of (9) and (10) is the key to the understanding of multiple orientations. The data tensors are invariant against any arbitrary permutation of indices and therefore have some very pronounced symmetry properties. For $M = 2$, the data tensors \mathcal{O} and \mathcal{T} are symmetric $N \times N$ matrices, but for higher M , we cannot rely on concepts from matrix algebra anymore. We therefore define the *space of fully symmetric M -tensors*³ as

$$\mathbb{R}_{\oplus}^{N \times \dots \times N} = \left\{ \mathcal{T} \in \mathbb{R}^{N \times \dots \times N} \mid (\mathcal{T})_{i_1 \dots i_M} = (\mathcal{T})_{P(i_1 \dots i_M)} \right\} \quad (11)$$

with $P(i_1 \dots i_M)$ denoting any arbitrary permutation of the indices $i_1 \dots i_M$.

Whereas the data tensors are fully symmetric, the orientation tensor $\mathcal{U} = \mathbf{u}_1 \otimes \dots \otimes \mathbf{u}_M$ is clearly not. But the symmetry of the left operand in some scalar product (like $\langle \mathcal{T}, \mathcal{U} \rangle$) always means that the value of the scalar product does not change if the same symmetry transformations are applied to the second operand (here: \mathcal{U}). Hence, if $\langle \mathcal{T}, \mathcal{U} \rangle = 0$, then $\langle \mathcal{T}, \mathcal{U}' \rangle = 0$ with \mathcal{U}' denoting any

³ Fully symmetric in order to allow the term symmetric also for invariance against *special* permutations only, for instance the exchange of indices 1 and 2.

arbitrary permutation of the order of orientations in the tensor product \mathcal{U} .⁴ As a consequence, the original tensor \mathcal{U} *cannot* be recovered uniquely and any linear combination of permuted tensors solves the problem.

However, it *is* possible to describe a set of M orientations with a unique order- M tensor. The key is symmetrization: among all possible orientation tensors \mathcal{U} which are orthogonal to \mathcal{T} (i.e. $\langle \mathcal{T}, \mathcal{U} \rangle = 0$), there is only a single fully symmetric one (up to a non-zero scale factor): the sum over all possible permutations with *equal* weights. This means that we have to estimate the orientation tensor subject to $\mathcal{U} \in \mathbb{R}_{\oplus}^{N \times \dots \times N}$ in order to obtain a unique solution.

All tensor scalar products can be converted to standard scalar products by stacking the tensor elements to form a long vector. But now, the symmetry properties of both operands call for a slightly modified version of vectorization. Our fully symmetric tensors have

$$k = \binom{N + M - 1}{M} \quad (12)$$

different elements (\equiv degrees of freedom, DOF). Therefore, the space $\mathbb{R}_{\oplus}^{N \times \dots \times N}$ can be mapped to \mathbb{R}^k . We now define

Definition 1. Let $\mathcal{A} \in \mathbb{R}_{\oplus}^{N \times \dots \times N}$ denote a fully symmetric tensor of order M . Then we define the mapping $\text{VecSymm}(\cdot) : \mathbb{R}_{\oplus}^{N \times \dots \times N} \rightarrow \mathbb{R}^k$ with k defined in (12) as stacking all independent elements under each other in some arbitrary but fixed order. Furthermore, we define $\text{VecSymmN}(\cdot) : \mathbb{R}_{\oplus}^{N \times \dots \times N} \rightarrow \mathbb{N}^k$ as counting the number of appearances (index permutations) of each element.

Note that the $\text{VecSymmN}(\cdot)$ operation only depends on the dimensionality of the argument, not on the entries. Thus, every element of $\mathbb{R}_{\oplus}^{N \times \dots \times N}$ produces the same $\text{VecSymmN}(\cdot)$ result. Applying these definitions to (9) and (10) now allows to generalize the single orientation constraint $\langle \mathbf{g}, \mathbf{u} \rangle = 0$ (gradient orthogonal to sought orientation) to $\langle \tilde{\mathbf{g}}, \tilde{\mathbf{u}} \rangle = 0$ with *mixed orientations gradient* vector

$$\text{(OOM:)} \quad \tilde{\mathbf{g}} = \text{VecSymm}(\mathcal{O}) \quad \text{resp.} \quad \text{(TOM:)} \quad \tilde{\mathbf{g}} = \text{VecSymm}(\mathcal{T}) \quad (13)$$

and *mixed orientation parameters* (MOP) vector

$$\tilde{\mathbf{u}} = \text{VecSymmN}(\mathcal{U}) \cdot \text{VecSymm}(\mathcal{U}) \quad (14)$$

with “ \cdot ” indicating element-by-element multiplication. The vectors $\tilde{\mathbf{g}}$ are the multiple orientations equivalent of the gradients. In analogy to single orientation estimation, we can therefore define the double [triple, M -] orientation structure tensor $\mathbf{S}^{(M)}$ as spatial integration (i.e. convolution with averaging filter f') over the outer product of $\tilde{\mathbf{g}}$ with itself

$$\mathbf{S}^{(M)} = f' * (\tilde{\mathbf{g}} \tilde{\mathbf{g}}^T) \quad (15)$$

⁴ Geometric interpretation: we can only estimate a *set* of M orientations and cannot identify them as “first”, “second”, “ M -th” orientation; they are interchangeable.

and the eigenvector of $\mathbf{S}^{(M)}$ corresponding to the smallest eigenvalue will then yield the MOP vector for M orientations. But unfortunately, we cannot claim that this already solves the multiple orientation estimation problem because

- a) the MOP vector has clearly too many DOF: k is a polynomial with leading term N^M , whereas M unit vectors only have $M(N - 1)$ DOF, and
- b) no general way exists in literature how to decompose it into its underlying orientations,⁵

i.e. what we have solved by now is just an intermediate step towards the sought set of orientation vectors.

2.3 Vector-Valued Signals

Based on early work by Di Zenzo [16] and Förstner [17], who first studied gradients of multi-band images, it is possible to derive multiple orientations structure tensors also for vector-valued signals $\mathbf{s}(\mathbf{x}) \in \mathbb{R}^P$ (for instance color images). Orientation estimation is also possible in such P -dimensional data, but notation gets much more complex. We therefore deferred discussion of multi-dimensional signals to this point where the generalized gradients $\tilde{\mathbf{g}}$ have become available.

In principle, for every derivative, we now have to choose between P signal bands. In the transparent model, each component of $\tilde{\mathbf{g}}$ is a single M -th order derivative; hence for general P , it gets vector-valued and for the structure tensor, we have to perform an additional contraction over this index, turning the outer product $\tilde{\mathbf{g}}\tilde{\mathbf{g}}^T$ into a matrix of scalar products.⁶ For instance, multi-dimensional TOM orientation estimation for $N = 2$ and $M = 2$ means

$$\mathbf{S}^{(2)} = f' * \begin{pmatrix} \langle \mathbf{s}_{xx}, \mathbf{s}_{xx} \rangle & \langle \mathbf{s}_{xx}, \mathbf{s}_{xy} \rangle & \langle \mathbf{s}_{xx}, \mathbf{s}_{yy} \rangle \\ \langle \mathbf{s}_{xx}, \mathbf{s}_{xy} \rangle & \langle \mathbf{s}_{xy}, \mathbf{s}_{xy} \rangle & \langle \mathbf{s}_{xy}, \mathbf{s}_{yy} \rangle \\ \langle \mathbf{s}_{xx}, \mathbf{s}_{yy} \rangle & \langle \mathbf{s}_{xy}, \mathbf{s}_{yy} \rangle & \langle \mathbf{s}_{yy}, \mathbf{s}_{yy} \rangle \end{pmatrix}. \quad (16)$$

Under the occlusion model (M first-order derivatives), every element of the mixed orientations gradient becomes a $P \times \dots \times P$ tensor (M factors) and computing the structure tensor means contraction over all M signal band indices. Fortunately, all elements are outer products which turns the structure tensor elements from products of $2M$ scalar values (for $P = 1$) to M scalar products of two P -vectors. Considering $N = 2$ and $M = 2$ for arbitrary P again yields $\tilde{\mathbf{g}} = (\mathbf{s}_x \otimes \mathbf{s}_x, \frac{1}{2}\mathbf{s}_x \otimes \mathbf{s}_y + \frac{1}{2}\mathbf{s}_y \otimes \mathbf{s}_x, \mathbf{s}_y \otimes \mathbf{s}_y)^T$ (symbolic notation!) and then

$$\mathbf{S}^{(2)} = f' * \begin{pmatrix} \langle \mathbf{s}_x, \mathbf{s}_x \rangle^2 & \langle \mathbf{s}_x, \mathbf{s}_x \rangle \langle \mathbf{s}_x, \mathbf{s}_y \rangle & \langle \mathbf{s}_x, \mathbf{s}_y \rangle^2 \\ \langle \mathbf{s}_x, \mathbf{s}_x \rangle \langle \mathbf{s}_x, \mathbf{s}_y \rangle & \frac{1}{2} \langle \mathbf{s}_x, \mathbf{s}_x \rangle \langle \mathbf{s}_y, \mathbf{s}_y \rangle + \frac{1}{2} \langle \mathbf{s}_x, \mathbf{s}_y \rangle^2 & \langle \mathbf{s}_x, \mathbf{s}_y \rangle \langle \mathbf{s}_y, \mathbf{s}_y \rangle \\ \langle \mathbf{s}_x, \mathbf{s}_y \rangle^2 & \langle \mathbf{s}_x, \mathbf{s}_y \rangle \langle \mathbf{s}_y, \mathbf{s}_y \rangle & \langle \mathbf{s}_y, \mathbf{s}_y \rangle^2 \end{pmatrix} \quad (17)$$

⁵ In principle, the MOP vector itself (i.e. without decomposition into underlying orientations) could be used for applications like texture classification or tracking. However, the distance between two vectors in an highly overparameterized space is clearly suboptimal without previous projection onto the space of “valid MOP vectors”.

⁶ As summation and convolution commute, we can alternatively sum the P structure tensors computed for each individual signal band.

(this equation was also derived in [15]). We can thus finally compute structure tensors for an arbitrary number M of orientations, either occludingly or transparently superposed, in arbitrary P -dimensional and N -variate signals.

3 Solving the Decomposition Problem

Once an estimate $\hat{\mathbf{u}}$ for the MOP vector is computed, we first reverse (14) by dividing each component by the corresponding number of permutations. Then the mapping itself can be reversed, thus producing an estimate $\hat{\mathcal{U}}$ which is a fully symmetric tensor, i.e. an element of $\mathbb{R}_{\oplus}^{N \times \dots \times N}$. However, this tensor is in general not an element of

$$\mathbb{R}_{\oplus}^{N \times \dots \times N} = \left\{ \sum_{P(i_1 \dots i_M)} \mathbf{u}_{i_1} \otimes \dots \otimes \mathbf{u}_{i_M} \mid \mathbf{u}_{i_1}, \dots, \mathbf{u}_{i_M} \in \mathbb{R}^N \setminus \{\mathbf{0}\} \right\}, \quad (18)$$

the space of *symmetrized* outer products, which we will call the space of *minimal fully symmetric tensors* from now on. Therefore, our estimate does not represent a valid set of M orientations in general. Going back to vector space, we see that the space of *valid MOP vectors* is a subset of \mathbb{R}^k , and only in tensor space, we have the means to define both spaces properly: any estimated tensor is a fully symmetric tensor (i.e. an element of $\mathbb{R}_{\oplus}^{N \times \dots \times N}$), but valid tensors have to be restricted to the subspace $\mathbb{R}_{\otimes}^{N \times \dots \times N}$. For single orientation estimation, this novel perspective on (multiple) orientation estimation coincides with the known definitions (order-1 tensors are vectors), but in general, only a tensor approach is suited to handle the symmetry constraints properly.

3.1 Multiple Orientation Estimation for Images

For bivariate images (i.e. $N = 2$), we find that $k = \binom{2+M-1}{M} = \binom{M+1}{M} = M + 1$. Subtracting 1 for undefined scale, we obtain M which is the same number as $M(N - 1)$ (DOF for M unit vectors in N -dimensional space) for $N = 2$. Therefore, the problem of overdetermined MOP vectors does not appear in images. This means that we have to qualify the last sentence of the previous paragraph: for images (and only for images!), the MOP vector is in fact a minimal description of the sought parameters. In images, derivatives are only possible with respect to two coordinates, say x and y . For instance, $M = 3$ yields

$$\begin{aligned} (\tilde{\mathbf{u}})_1 &= (\mathbf{u}_1)_x (\mathbf{u}_2)_x (\mathbf{u}_3)_x \\ (\tilde{\mathbf{u}})_2 &= (\mathbf{u}_1)_x (\mathbf{u}_2)_x (\mathbf{u}_3)_y + (\mathbf{u}_1)_x (\mathbf{u}_2)_y (\mathbf{u}_3)_x + (\mathbf{u}_1)_y (\mathbf{u}_2)_x (\mathbf{u}_3)_x \\ (\tilde{\mathbf{u}})_3 &= (\mathbf{u}_1)_x (\mathbf{u}_2)_y (\mathbf{u}_3)_y + (\mathbf{u}_1)_y (\mathbf{u}_2)_x (\mathbf{u}_3)_y + (\mathbf{u}_1)_y (\mathbf{u}_2)_y (\mathbf{u}_3)_x \\ (\tilde{\mathbf{u}})_4 &= (\mathbf{u}_1)_y (\mathbf{u}_2)_y (\mathbf{u}_3)_y ; \end{aligned}$$

generalization to arbitrary M is straightforward. Every orientation vector consists of two unknowns $(\mathbf{u}_i)_x$ and $(\mathbf{u}_i)_y$ with the constraint $((\mathbf{u}_i)_x)^2 + ((\mathbf{u}_i)_y)^2 = 1$ (it must be a unit vector). However, this problem can be reformulated as an

unconstrained problem easily: If $(\tilde{\mathbf{u}})_1 = 0$, then at least one of the sought orientations is $(0, 1)^T$. Without loss of generality, we therefore can define $(\mathbf{u}_M)_x = 0$ and $(\mathbf{u}_M)_y = 1$, thus reducing the degree of the problem by 1. Otherwise, we divide by $(\tilde{\mathbf{u}})_1$ and obtain the equation system

$$\begin{aligned} p_1 &= x_1 + x_2 + x_3 \\ p_2 &= x_1x_2 + x_1x_3 + x_2x_3 \\ p_3 &= x_1x_2x_3 \end{aligned}$$

with given values $p_i := \frac{(\tilde{\mathbf{u}})_{i+1}}{(\tilde{\mathbf{u}})_1}$ and the new unknowns $x_i := \frac{(\mathbf{u}_i)_y}{(\mathbf{u}_i)_x}$. The set of M values for x_i (which can be interpreted as slope of the orientation vectors) are easily found as roots of the polynomial

$$\sum_{i=0}^M (-1)^i (\tilde{\mathbf{u}})_{i+1} x^{M-i} = 0. \tag{19}$$

By combining the vertical orientation vectors with the normalized version of all $(1, x_i)^T$ vectors, we successfully solved the multiple orientations estimation problem for images.

3.2 Double Orientation Estimation in Multivariate Signals

For double orientation estimation, all tensors can be interpreted as matrices. Matrix algebra offers a convenient interpretation of the difference between fully symmetric tensors $(\mathbb{R}_{\oplus}^{N \times N})$ and its subset $\mathbb{R}_{\otimes}^{N \times N}$. While the first space is the space of symmetric $N \times N$ matrices, the latter space is the space of matrices formed by $\mathbf{u}_1 \otimes \mathbf{u}_2 + \mathbf{u}_2 \otimes \mathbf{u}_1$, i.e. the space of symmetric *rank-2* matrices.

This allows to define a very simple strategy for double-orientation estimation in general N -variate signals. We estimate the MOP vector $\tilde{\mathbf{u}}$ and map it to the space of fully symmetric tensors (here: symmetric matrices), taking care not to forget the division by the permutation count, see (14). Let \mathbf{U} denote the result of this operation; we now have to find the two unit vectors \mathbf{u}_1 and \mathbf{u}_2 which fulfill

$$\mathbf{U} = c(\mathbf{u}_1 \otimes \mathbf{u}_2 + \mathbf{u}_2 \otimes \mathbf{u}_1) \tag{20}$$

for some scaling factor c . From

$$\mathbf{U}\mathbf{u}_1 = c(c' \mathbf{u}_1 + \mathbf{u}_2) \quad \text{and} \quad \mathbf{U}\mathbf{u}_2 = c(\mathbf{u}_1 + c' \mathbf{u}_2) \tag{21}$$

(with $c' = \langle \mathbf{u}_1, \mathbf{u}_2 \rangle$ denoting the cosine of the angle between \mathbf{u}_1 and \mathbf{u}_2) follows

$$\mathbf{U}(\mathbf{u}_1 + \mathbf{u}_2) = c(c' + 1)(\mathbf{u}_1 + \mathbf{u}_2) \quad \text{and} \quad \mathbf{U}(\mathbf{u}_1 - \mathbf{u}_2) = c(c' - 1)(\mathbf{u}_1 - \mathbf{u}_2), \tag{22}$$

i.e. $\mathbf{u}_1 - \mathbf{u}_2$ and $\mathbf{u}_1 + \mathbf{u}_2$ are eigenvectors of \mathbf{U} . Being a rank-2 matrix, \mathbf{U} only has two non-zero eigenvalues and because $|\langle \mathbf{u}_1, \mathbf{u}_2 \rangle| < 1$ for $\mathbf{u}_1 \neq \mathbf{u}_2$, one of them is positive and one is negative. Let λ_+, λ_- denote the eigenvalues and $\mathbf{x}_+,$

\mathbf{x}_- the corresponding eigenvectors of \mathbf{U} (+ for positive eigenvalue, - for negative eigenvalue). Then the sought orientations can be found by normalizing

$$\mathbf{u}_1 = \sqrt{\lambda_+} \mathbf{x}_+ + \sqrt{-\lambda_-} \mathbf{x}_- \quad \text{and} \quad \mathbf{u}_2 = \sqrt{\lambda_+} \mathbf{x}_+ - \sqrt{-\lambda_-} \mathbf{x}_-. \quad (23)$$

Given some noisy estimate for \mathbf{U} , we can apply exactly the same strategy, now silently ignoring all intermediate eigenvalues which are close to zero (instead of being exactly zero); this yields the closest minimal fully symmetric matrix in terms of the Frobenius norm. Our scheme closely resembles to a method proposed by Shizawa and Mase for trivariate signals ($N = 3$) in [10]. The difference to our scheme is that \mathbf{U} was corrected to the space of 3×3 -matrices having rank 2 by subtracting λ_2 times the identity matrix with λ_2 denoting the (only) intermediate eigenvalue. In contrast to our scheme, this method cannot be generalized to $N > 3$ because it only works for a single intermediate eigenvalue.

Summarizing this section, we have presented multiple orientation estimation schemes for (a) $N = 2$ and arbitrary M and (b) $M = 2$ and arbitrary N . Again we stress that the decomposition schemes can be applied for both models (occluded and transparent) and for either grey value or multispectral data.

4 Experiments

We tested our algorithms on synthetic and real data. Synthetic data allow a thorough examination of the performance of an algorithm with known ground truth (under some assumed model), while testing with real data gives evidence that this modelling is accurate (at least for specific situations). Both types complement one another.

For multiple orientation estimation, the amount of parameters which can be modified is huge: the number M of superposed signals, the basic signals themselves, their orientations, their respective weight functions in the combination process, the filters used for computing the discrete derivatives, the filter used for spatial integration of the structure tensors, and the level and type of added noise. In our opinion, a thorough examination and testing of an algorithm is only possible with synthetic data and a graphical user interface (GUI) which allows easy modification of the individual input parameters and immediate feedback on the consequences it has for the estimate. Fig. 1 shows our GUI tool which we used to “explore the parameter space”. It is enclosed on the electronic version of the proceedings and it is also available for download at www.lfb.rwth-aachen.de.

We found that orientation estimation under the models discussed above behaves rather robust under added noise. Figs. 2 and 3 show two examples for the combination of two transparent resp. three occluding orientations, both for added Gaussian noise with SNR of 6 dB. Thick lines: true orientations; darker thin lines: estimated orientations.

For our next experiment, we used the image of a house (fig. 4) and tested the ratio s between smallest and second-smallest eigenvalue of $\mathbf{S}^{(M)}$ against some predefined thresholds. This allows a hierarchical testing of orientedness: if the texture (norm of image gradient) is high enough, then set $M = 1$ and

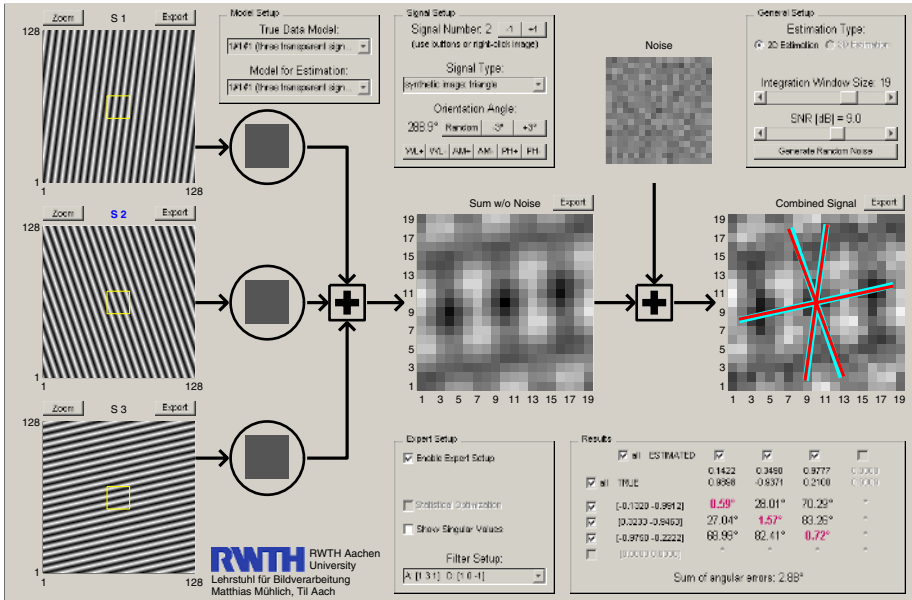


Fig. 1. The GUI which we use for experimentation with synthetic data

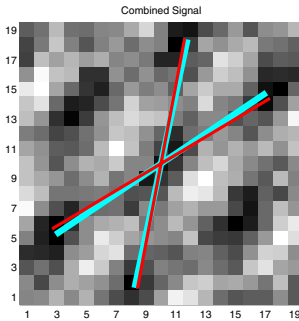


Fig. 2. Estimation of 2 transparent orientations (SNR = 6 dB)

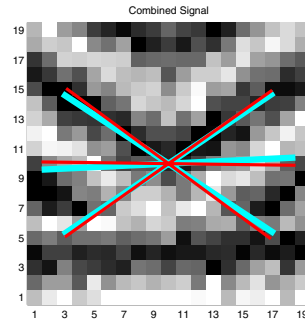


Fig. 3. Estimation of 3 occluding orientations (SNR = 6 dB)

1. compute structure tensor $\mathbf{S}^{(M)}$ and significance value s
2. if s is lower than some threshold c_M , then compute orientation vectors for M orientations model
3. otherwise increase M by one and go to first step (provided that M is smaller or equal than some maximum value M_{\max}).

Applying this scheme with $M_{\max} = 3$, we obtain a segmentation of the image into areas of 1, 2, or 3 orientations, plus two classes for “not enough texture” (i.e. image more or less constant) and “not oriented” (which is a reasonable



Fig. 4. An image of a house. Different textures can be characterized by a different number of local orientations.



Fig. 5. Number of orientations (0-3) encoded in four different gray levels. The regions in white were rejected by our orientation model.

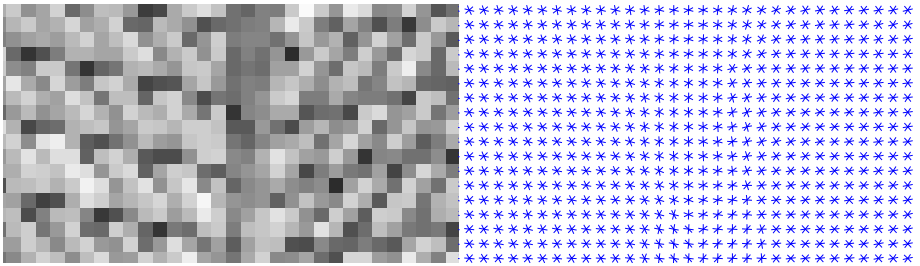


Fig. 6. Image (left) and estimated orientations (right) for each pixel at the part of the image where the two roof parts meet. We can see that the roof tiles are modelled well with *three* local orientations, and the two textures / roofs lead to two different sets of orientation estimates.

interpretation for bivariate data not fulfilling our model with $M \leq 3$). Fig. 5 shows the segmentation result which is based *only* on the number of found orientations. Also note that the image content (a house) is clearly visible in an image with only 5 different values; this demonstrates the importance of (multiple) local orientations as low level images (resp. texture) features.

For each of the regions labelled as single-, double- or triple-oriented, we also obtain the corresponding orientation vectors. Fig. 6 shows the part of the image where the two roof parts meet and it is clearly visible that the three estimated orientations do not vary much within the same roof, but are considerably different in both halves of the image (the ‘stars’ indicating the three orientations roughly look mirrored), thus allowing further segmentations within the regions having the same number of local orientations.

Another important application for the low level image feature “local orientation(s)” is the definition of invariance properties. For instance, the angle between two orientations defines a measure which is invariant to rotation. Therefore, the

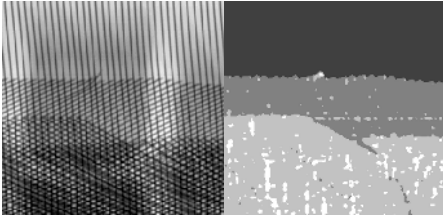


Fig. 7. Left: An x-ray image of metal gratings in a rubber product. Right: number of local orientations for each pixel.

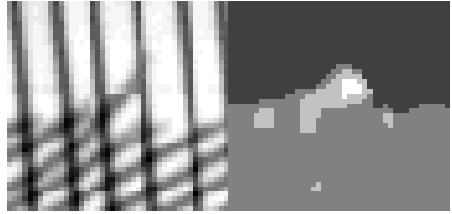


Fig. 8. Left: enlarged part of fig. 7. Right: number of local orientations again; the defect is easily visible.

theory of multiple orientation estimation is also important for the search of corresponding regions.

A third set of experiments was carried out on x-ray images within an industrial inspection application. Fig. 7 shows an x-ray image of metal gratings in a rubber product. The superposition of such gratings gives rise to single-, double- or triple-oriented areas. This means that the theory presented in our paper can be used to detect the number of gratings (i.e. the number of orientations) and their respective orientations vectors.

More specifically, it also allows to detect defects where a metal wire is not aligned properly within the grating or extends beyond the edge. Fig. 8 shows an enlarged part of fig. 7 where such a defect is visible as model violation and therefore increase of orientation number. In the segmentation image, the defect is clearly visible as a blob of triple-oriented and non-oriented textures.

5 Summary and Conclusions

In this paper, we presented a theory for modelling textures composed from multiple dominant orientations, thus extending the well-known structure tensor framework to a unified mathematical model for M orientations in P -dimensional and N -variate signals $\mathbf{s}(\mathbf{x})$. Generalization of the signal gradient to multiple directional derivatives leads to tensor-valued entities, and depending on the assumed signal model, this generalization can be done either under the occluding orientations model (OOM) or under the transparent orientations model (TOM). (One can also imagine mixed or intermediate forms which could be a topic for future research.) After the discussion of the two ways for generating data tensors, we have identified a suitable mathematical representation for a set of M orientations, namely the space of *minimal fully symmetric tensors* $\mathbb{R}_{\otimes}^{N \times \dots \times N}$. We emphasize that this tensor representation is superior to other approaches relying on vectorization. The mixed orientation parameters (MOP) vector – which we derived for general M , N and P – can be a highly overparameterized representation of orientations. Based on this better mathematical understanding of multiple orientation estimation, we presented algorithms for multiple orientation estimation in images (the only case where the MOP vector is *not* overparameterized) and for double

orientation estimation in general N -variate data. Both algorithms are applicable also to vector-valued data, for instance color images.

In the experimental part, we successfully applied multiple orientation estimation (in contrast to previous papers: beyond double orientations) to both synthetic and real data. Especially for image data, we showed that estimation of superimposed orientations (here: with $M = 0, \dots, 3$) provides new and highly useful low level image features which appear perfectly suited for various inspection, tracking or segmentation problems.

References

1. Förstner, W.: A feature based corresponding algorithm for image matching. *Intl. Arch. of Photogrammetry and Remote Sensing* **26** (1986) 150–166
2. Bigün, J., Granlund, G.H.: Optimal orientation detection of linear symmetry. In: *Proc. ICCV87, London, UK* (1987) 433–438
3. Mester, R.: The generalization, optimization and information-theoretic justification of filter-based and autocovariance based motion estimation. In: *IEEE International Conference on Image Processing, Barcelona, Spain* (2003)
4. Granlund, G., Knutsson, H.: *Signal Processing for Computer Vision*. Kluwer (1995)
5. Eberly, D., Gardner, R., Morse, B., Pizer, S., Scharlach, C.: Ridges for image analysis. Technical Report TR93-055, Dep. of Computer Science, Univ. of North Carolina, Chapel Hill, NC, USA (1993)
6. Jacob, M., Unser, M.: Design of steerable filters for feature detection using canny-like criteria. *IEEE Trans. PAMI* **26** (2004) 1007–1019
7. Felsberg, M., Granlund, G.: POI detection using channel clustering and the 2d energy tensor. In: *Proc. DAGM04, Springer Verlag: LNCS 3175* (2004) 103–110
8. Köthe, U.: Integrated edge and junction detection with the boundary tensor. In: *Proc. ICCV03. Volume 1, Nice, France* (2003) 424–431
9. Köthe, U., Felsberg, M.: Riesz-transforms versus derivatives: On the relationship between the boundary tensor and the energy tensor. In: *Proc. Scale Space and PDE Methods in Computer Vision, Springer Verlag: LNCS 3459* (2005) 179–191
10. Shizawa, M., Mase, K.: Simultaneous multiple optical flow estimation. In: *Proc. Int. Conf. Pattern Recognition*. (1990) 274–278
11. Shizawa, M., Mase, K.: A unified computational theory for motion transparency and motion boundaries based on eigenenergy analysis. In: *Proc. CVPR 1991*.(1991) 289–295
12. Shizawa, M., Iso, T.: Direct representation and detecting of multi-scale, multi-orientation fields using local differentiation filters. In: *Proc. CVPR 1993*. (1993) 508–514
13. Aach, T., Stuke, I., Mota, C., Barth, E.: Estimation of multiple local orientations in image signals. In: *Proc. ICASSP04, Montreal, Canada* (2004) 553–556
14. Mota, C., Aach, T., Stuke, I., Barth, E.: Estimation of multiple orientations in multi-dimensional signals. In: *Proc. ICIP04, Singapore* (2004) 2665–2668
15. Mota, C., Stuke, I., Aach, T., Barth, E.: Estimation of multiple orientations at corners and junctions. In: *Proc. DAGM04, Tübingen, Germany, Springer Verlag: LNCS 3175* (2004) 163–170
16. Di Zenzo, S.: A note on the gradient of a multi-image. *Computer Vision, Graphics, and Image Processing* **33** (1986) 116–125
17. Förstner, W., Gülch, E.: A fast operator for detection and precise location of distinct points, corners and centres of circular patterns. In: *Proc. ISPRS Intercomm. Workshop*. (1987) 281–305