14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

TOWARDS MULTIPLE-ORIENTATION BASED TENSOR INVARIANTS FOR OBJECT TRACKING

Nicolaj C. Stache, Thomas H. Stehle, Matthias Mühlich, and Til Aach

Institute of Imaging and Computer Vision, RWTH Aachen University
D-52056 Aachen, Germany
phone: + (49) 241 8027860, fax: + (49) 241 8022200
email: {stache,stehle,muehlich,aach}@lfb.rwth-aachen.de, web: www.lfb.rwth-aachen.de

ABSTRACT

We derive a new scale- and rotation-invariant feature for characterizing local neighbourhoods in images, which is applicable in tasks such as tracking. Our approach is motivated by the estimation of optical flow. Its least-squares estimate requires the inversion of a symmetric and positive semi-definite 2×2-tensor, which is computed from spatial image derivatives. Only if one eigenvalue of the tensor vanishes, this tensor describes the local neighbourhood in terms of orientation. Estimating optical flow, however, requires that this tensor be regular, i.e., that both its eigenvalues do not vanish. This indicates that the local region contains more than one orientation.

Double-orientation neighbourhoods (like X junctions or corners) are especially suited for tracking or optical flow estimation, but the two underlying orientations cannot be extracted from the standard structure tensor. Therefore, we extend this tensor such that it can characterize double-orientation neighbourhoods. From this extended tensor, we derive a rotation- and scale-invariant feature which describes the orientation structure of the local regions, and analyze its performance.

1. INTRODUCTION

In this paper, we derive a rotation- and scale-invariant feature for the description of local image neighbourhoods, which may be applied in tasks such as object tracking. The feature is based on the analysis of the orientation structure of the image signal in local regions by tensors, which are formed from spatial derivatives. Though thus aimed at motion analysis by feature matching rather than differential estimation of motion, let us motivate our approach by the differential estimation of optical flow. Let $f(\mathbf{x},t)$, with $f : \mathbb{R}^3 \to \mathbb{R}$, denote a grey-level image sequence, where $\mathbf{x} = (x,y)^T \in \mathbb{R}^2$ is the spatial coordinate vector, and $t \in \mathbb{R}$ is time (additionally, we assume that f is differentiable). The fundamental constraint for estimating the optical flow field $\mathbf{w}(\mathbf{x},t) = (w_x(\mathbf{x},t), w_y(\mathbf{x},t))$, $\mathbf{w} : \mathbb{R}^3 \to \mathbb{R}^2$, is formed by setting the total temporal derivative of f to zero [6, 8, 14]:

$$\frac{d}{dt} f(\mathbf{x},t) = (\nabla f)^T \cdot \mathbf{w} + f_t = 0 \ . \tag{1}$$

Here, $\nabla f = (f_x, f_y)$ is the spatial gradient of the image signal, while f_x, f_y and f_t are the partial derivatives of $f(\mathbf{x},t)$ with respect to x, y and t. Practically, the flow field is assumed to be constant within a small spatial region $\Omega \subset \mathbb{R}^2$. The least-squares solution for $\mathbf{w}(\mathbf{x},t)$ $\forall \mathbf{x} \in \Omega$ then obeys

$$\mathbf{J}\mathbf{w} = \mathbf{b} \Rightarrow \mathbf{w} = \mathbf{J}^{-1}\mathbf{b} \ , \tag{2}$$

where $\mathbf{b}$ is a twodimensional vector calculated from the partial image derivatives. The tensor $\mathbf{J}$ is a symmetric and positive semi-definite 2×2-matrix, which is calculated from the spatial image gradient by

$$\mathbf{J} = \int_\Omega (\nabla f)(\nabla f)^T d\Omega = \int_\Omega \begin{bmatrix} f_x^2 & f_x f_y \\ f_x f_y & f_y^2 \end{bmatrix} d\Omega \ . \tag{3}$$

Inversion of $\mathbf{J}$ is possible if it is regular, i.e. if both its eigenvalues do not vanish. This condition can also be interpreted as follows: the image signal within Ω must *not* be characterized by a single orientation – the matrix $\mathbf{J}$ is the very same entity which is also known as *structure tensor*, and which can be used to estimate the local spatial orientation as the eigenvector of $\mathbf{J}$ corresponding to the smallest (ideally: zero) eigenvalue [3, 7, 2].

If one eigenvalue is close to zero, the signal is oriented and we can determine this orientation – but a unique optical flow estimation is prevented by the aperture problem. On the other hand, if no eigenvalues are close to zero, we can estimate optical flow, but the corresponding eigenvectors do not represent orientations, e.g. if the signal is a superposition of two (single-)oriented signals.

We therefore describe an appropriately extended tensor which can capture two orientations of, e.g., a corner. The eigensystem analysis of this tensor, however, does not directly yield the sought orientations, rather, it provides a so-called mixed-orientation parameter (MOP) vector, which implicitly encodes both sought orientations, and is thus not rotation invariant. Instead of decomposing the MOP vector into the sought orientations, we discuss here how to derive a feature from it which unambiguously characterizes the local orientation structure, and is invariant to rotation as well as to certain scalings of the coordinate axes and intensity. Most of our discussion will, for ease of notation, be developed for bivariate image data. We will also, however, show how to extend the framework towards higher-variate data, such as tomograms.

2. DOUBLE ORIENTATIONS

2.1 Bivariate Images

Let us now consider a bivariate image $f(\mathbf{x})$, with $f : \mathbb{R}^2 \to \mathbb{R}$, to be additively composed within a local region Ω from two oriented subimages by [12, 13, 1]

$$f(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x}), \quad f_1(\mathbf{x}),\, f_2(\mathbf{x}) \ : \ \mathbb{R}^2 \to \mathbb{R} \ . \tag{4}$$

Subimage $f_1(\mathbf{x})$ is oriented along θ with the orientation vector $\mathbf{u} = (\cos\theta, \sin\theta)^T = (u_x, u_y)^T$, while $f_2(\mathbf{x})$ is oriented

along γ with the orientation vector $\mathbf{v} = (\cos\gamma, \sin\gamma)^T = (v_x, v_y)^T$. Therefore, $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ obey

$$\alpha(\theta) f_1(\mathbf{x}) = \mathbf{u}^T \cdot \nabla f_1(\mathbf{x}) = 0 \quad \forall \mathbf{x} \in \Omega \tag{5}$$

$$\alpha(\gamma) f_2(\mathbf{x}) = \mathbf{v}^T \cdot \nabla f_2(\mathbf{x}) = 0 \quad \forall \mathbf{x} \in \Omega \tag{6}$$

where $\alpha(\theta)$ denotes the directional derivative operator along θ. Alternatively, we consider the superposition of subimages which occlude each other [10]. In some part $\Omega_1 \subset \Omega$, we have $f(\mathbf{x}) = f_1(\mathbf{x})$, while in its complement Ω_2, $f(\mathbf{x}) = f_2(\mathbf{x})$. Our model then is

$$f(\mathbf{x}) = \begin{cases} f_1(\mathbf{x}) & \forall \mathbf{x} \in \Omega_1 \\ f_2(\mathbf{x}) & \forall \mathbf{x} \in \Omega_2 \end{cases} \tag{7}$$

with $\frac{\partial f_1(\mathbf{x})}{\partial \mathbf{u}} = 0 \ \forall \mathbf{x} \in \Omega_1$, $\frac{\partial f_2(\mathbf{x})}{\partial \mathbf{v}} = 0 \ \forall \mathbf{x} \in \Omega_2$ and $\Omega_1 \cup \Omega_2 = \Omega$, $\Omega_1 \cap \Omega_2 = \emptyset$. In both cases, applying the operators $\alpha(\theta)$ and $\alpha(\gamma)$ sequentially to the composite image $f(\mathbf{x})$ yields the constraint

$$\alpha(\theta)\alpha(\gamma) f(\mathbf{x}) = \frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{u} \partial \mathbf{v}} = 0 \quad \forall \mathbf{x} \in \Omega \ . \tag{8}$$

(We neglect here that, in the occluding case, this constraint may be violated on the border $\partial\Omega_{12}$ between the subregions Ω_1 and Ω_2). We rewrite this constraint as the inner product

$$\mathbf{a}^T \mathbf{d} f(\mathbf{x}) = 0 \ \forall \mathbf{x} \in \Omega \tag{9}$$

where the three-dimensional vector $\mathbf{a}$ is given by

$$\begin{aligned} \mathbf{a}^T &= (u_x v_x, \ u_x v_y + u_y v_x, \ u_y v_y) \\ &= (\cos\theta\cos\gamma, \ \sin(\theta+\gamma), \ \sin\theta\sin\gamma) = (a, \ b, \ c) \end{aligned} \tag{10}$$

and

$$\mathbf{d} f(\mathbf{x}) = (f_{xx}, f_{xy}, f_{yy})^T \ . \tag{11}$$

The components of $\mathbf{a}$ are the so-called mixed orientation parameters (MOP) resulting from the concatenation of two directional derivatives. The least-squares solution for the MOP vector $\mathbf{a}$ then minimizes

$$Q(\mathbf{a}) = \int_\Omega [\mathbf{a}^T \mathbf{d} f]^2 d\Omega = \mathbf{a}^T \mathbf{T} \mathbf{a} = 0, \quad \mathbf{a}^T \mathbf{a} > 0 \tag{12}$$

where $\mathbf{T}$ is the 3×3-tensor

$$\begin{aligned} \mathbf{T} &= \int_\Omega (\mathbf{d} f)(\mathbf{d} f)^T d\Omega \\ &= \int_\Omega \begin{bmatrix} f_{xx}^2 & f_{xx} f_{xy} & f_{xx} f_{yy} \\ f_{xx} f_{xy} & f_{xy}^2 & f_{xy} f_{yy} \\ f_{xx} f_{yy} & f_{xy} f_{yy} & f_{yy}^2 \end{bmatrix} d\Omega \ . \end{aligned} \tag{13}$$

Minimizing $Q(\mathbf{a})$ subject to $\mathbf{a}^T \mathbf{a} > 0$ implies that $\mathbf{a}$ is the eigenvector of the $\mathbf{T}$ corresponding to its smallest eigenvalue λ_3:

$$\mathbf{T}\mathbf{a} = \lambda_3 \mathbf{a}, \quad \mathbf{a}^T \mathbf{a} = 1 \ . \tag{14}$$

The eigensystem analysis determines only the direction of $\mathbf{a}$. Setting $\mathbf{a}^T \mathbf{a} = 1$ as done above implies that the estimate is only known up to an unknown scaling factor R. In ideal double orientation neighbourhoods in the sense of Eqs. (4) or (7), the lowest eigenvalue λ_3 vanishes, i.e. $\mathrm{rank}(\mathbf{T}) = 2$.

The MOP vector implicitly encodes the orientation vectors $\mathbf{u}$ and $\mathbf{v}$. Methods to decompose $\mathbf{a}$ into $\mathbf{u}$ and $\mathbf{v}$ are described in [12, 13, 1, 9]. Here, we seek to extract a rotation-invariant feature from $\mathbf{a}$ without decomposing it first. Towards this end, we consider its degrees of freedom (DoF): As discussed above, $\mathbf{a}$ obeys the homogeneous equation $\mathbf{T}\mathbf{a} = \lambda_3 \mathbf{a} = \mathbf{0}$, and can only be determined up to scale and sign. The MOP vector therefore is an element of a projective space, where two vectors are equivalent when they differ only in norm and sign [5]. We may therefore constrain $\mathbf{a}$ to length one, doing so reduces its DoF from three to two. This number is equal to the number of parameters, viz. θ and γ, specifying two orientations in an image. Deriving the rotation invariant feature from $\mathbf{a}$ implies the loss of one DoF. Our feature therefore will only exhibit one DoF, i.e., it is scalar. Intuitively, the sought feature is then given by the difference angle β between the orientations,[1] or by functions of β, such as $|\cos\beta|$. From the MOP vector entries a, b, c defined in Eq. (10), the latter can be directly computed according to

$$|\cos\beta| = \frac{|a+c|}{\sqrt{(a-c)^2 + b^2}} \ . \tag{15}$$

Note that this result remains unchanged when scaling $\mathbf{a}$.

In the following, we generalize this approach to tri- and higher-variate input data, and prove that indeed all scalar invariants of $f(\mathbf{x})$ encoded in the MOP vector $\mathbf{a}$ are generated by the angle β.

2.2 Tri- and Higher-Variate Data

Let $f(\mathbf{x})$, $\mathbf{x} = (x_1, x_2, \ldots, x_p)^T$, $f : \mathbb{R}^p \to \mathbb{R}$ now denote a p-variate mapping. Within the local region $\Omega \subset \mathbb{R}^p$, let $f(\mathbf{x})$ consist of the additive or occluding superposition of two p-variate signals $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$, each of which be oriented along a line. Hence,

$$\frac{\partial}{\partial \mathbf{u}} f_1(\mathbf{x}) = \frac{\partial}{\partial \mathbf{v}} f_2(\mathbf{x}) = 0 \ . \tag{16}$$

Constraint (8) then expands to

$$\begin{aligned} \frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{u} \partial \mathbf{v}} &= \left(\sum_{i=1}^p u_i \frac{\partial}{\partial_i} \cdot \sum_{j=1}^p v_j \frac{\partial}{\partial_j} \right) f(\mathbf{x}) \\ &= \sum_{i=1}^p \sum_{j=1}^i a_{ij} f_{ij} = 0 \end{aligned} \tag{17}$$

where f_{ij} is the partial derivative of f with respect to x_i and x_j, while u_i, v_j, $i, j = 1, \ldots, p$ are the components of $\mathbf{u}$ and $\mathbf{v}$. The MOPs a_{ij} are given by

$$a_{ij} = \begin{cases} u_j v_j \text{ for } i = j \\ u_i v_j + u_j v_i \text{ else} \end{cases} \ . \tag{18}$$

Gathering the a_{ij} into the MOP vector $\mathbf{a} = [a_{ij}]_{j \le i}^T$ and with the vector $\mathbf{d} f = [f_{ij}]_{j \le i}^T$ of second derivatives, we obtain from

[1] More precisely, it is the pair $(\beta, 180^\circ - \beta)$. Note that both angles share the same absolute value of the cosine; a negative sign correponds to an angle greater than 90°, while $\cos\beta \ge 0$ means $\beta \le 90^\circ$. Therefore, restricting $\cos\beta$ to non-negative values by taking the absolute value automatically chooses the *smaller* angle of the pair $(\beta, 180^\circ - \beta)$.

Eq. (17)

$$\mathbf{a}^T \mathbf{d}f(\mathbf{x}) = 0 \;\; \forall \mathbf{x} \in \Omega \tag{19}$$

which is structurally similar to Eq. (9). The sum in Eq. (17) consists of $k = p(p+1)/2$ components. With the symmetric $k \times k$-tensor

$$\mathbf{T} = \int_\Omega (\mathbf{d}f)(\mathbf{d}f)^T d\Omega \tag{20}$$

the MOP vector $\mathbf{a}$ satisfies $\mathbf{a}^T\mathbf{T}\mathbf{a} = 0$, $\mathbf{a}^T\mathbf{a} = 1$. As above, $\mathbf{a}$ is a homogeneous vector, we thus may set $\mathbf{a}^T\mathbf{a} = 1$.

Let us now consider another tensor $\mathbf{A}$ which is formed from the (as yet unknown) orientation vectors by $\mathbf{A} = \frac{1}{2}(\mathbf{u}\mathbf{v}^T + \mathbf{v}\mathbf{u}^T)$. The rank of $\mathbf{A}$ is two, thus, $\mathbf{A}$ has two non-vanishing eigenvalues λ_+ and λ_-. These eigenvalues are easily derived to

$$\lambda_+ = \frac{1}{2}[1 + \mathbf{u}^T\mathbf{v}] = \cos^2\frac{\beta}{2} > 0 \tag{21}$$

$$\lambda_- = \frac{1}{2}[\mathbf{u}^T\mathbf{v} - 1] = -\sin^2\frac{\beta}{2} < 0 \;\; , \tag{22}$$

where we have used that the inner product of the unit orientation vectors $\mathbf{u}$ and $\mathbf{v}$ yields the cosine of the difference angle β, i.e. $\mathbf{u}^T\mathbf{v} = \cos\beta$. Therefore,

$$\lambda_+ + \lambda_- = \cos\beta \tag{23}$$

$$\lambda_+ - \lambda_- = 1 \;\; . \tag{24}$$

With Eq. (18), this tensor can also be computed from the MOPs by setting $[A]_{jj} = a_{jj}$, and $[A]_{ij} = \frac{1}{2}a_{ij}$, $i \neq j$. We are now able to state which scalar invariant features of $f(\mathbf{x})$ are encoded in the MOP vector $\mathbf{a}$: The rotation of the content of the image $f(\mathbf{x})$ in Ω causes a rotation of the vectors $\mathbf{u}$ and $\mathbf{v}$. This can be described as a multiplication with the rotation matrix $\mathbf{M}$

$$\hat{\mathbf{u}} = \mathbf{M}\cdot\mathbf{u}, \qquad \hat{\mathbf{v}} = \mathbf{M}\cdot\mathbf{v} \tag{25}$$

and yields a rotated tensor

$$\begin{aligned}\hat{\mathbf{A}} &= \frac{1}{2}\left(\mathbf{M}\mathbf{u}\cdot(\mathbf{M}\mathbf{v})^T + \mathbf{M}\mathbf{v}\cdot(\mathbf{M}\mathbf{u})^T\right) \\ &= \mathbf{M}\frac{1}{2}\left(\mathbf{u}\mathbf{v}^T + \mathbf{v}\mathbf{u}^T\right)\mathbf{M}^T = \mathbf{M}\mathbf{A}\mathbf{M}^T \;\; .\end{aligned} \tag{26}$$

Since rotation is an orthogonal similarity transform, the eigenvalues of $\hat{\mathbf{A}}$ are identical to those of $\mathbf{A}$. All scalar invariants under local orthogonal similarity transforms are thus generated by λ_+ and λ_- or, from Eq. (23), by the angle β between the orientations.

The algorithm for calculating the feature $|\cos\beta|$ is therefore as follows:

- Calculate $\mathbf{T}$ from the second-order image derivatives over Ω.
- Estimate the MOP vector as the eigenvector belonging to the lowest eigenvalue λ_3 of $\mathbf{T}$.
- Form the tensor $\mathbf{A}$ from the entries of $\mathbf{a}$.
- Calculate the non-zero eigenvalues $\lambda_+ > 0$ and $\lambda_- < 0$ of $\mathbf{A}$, and confirm that all other eigenvalues, if any, are (close to) zero.
- Calculate $\cos\beta$ from the eigenvalues: Since the vector $\mathbf{a}$ and hence the tensor $\mathbf{A}$ are only known up to a scaling factor R, Eqn. (23) and (24) change to $\lambda_+ + \lambda_- = R\cos\beta$ and $\lambda_+ - \lambda_- = R$, yielding for $\cos\beta$

$$\cos\beta = \frac{\lambda_+ + \lambda_-}{\lambda_+ - \lambda_-} \;\; . \tag{27}$$

As both numerator and denominator are positive, the same holds for $\cos\beta$ and it follows $\beta \leq 90°$ (see also footnote 1; the identification of a "positive" and a "negative" eigenvalue here is equivalent to taking the absolute value in Eq. (15))

2.3 Scaling Invariances

We now briefly discuss invariances of the rotation-invariant measure $|\cos\beta|$ as calculated from Eq. (27) with respect to intensity scaling and scaling of the coordinate axes. Let $\hat{f} = c_g f + c_o$, $c_g, c_o \in \mathbb{R}$, $c_g \neq 0$, denote the intensity-scaled version of f. In the tensor $\mathbf{T}$, this results in a multiplication of each entry by c_g^2. The eigenvector $\mathbf{a}$, though, remains unchanged, and so does, consequently, $|\cos\beta|$. Similarly, scaling each component x_i, $i = 1, \ldots, p$, of the space vector $\mathbf{x}$ by the same factor $c_s \in \mathbb{R}$, $c_s \neq 0$, yields $\hat{f}(\mathbf{x}) = f(c_s \cdot \mathbf{x})$. Each entry of $\mathbf{T}$ is then multiplied by the same factor c_s^4, again leaving the MOP vector $\mathbf{a}$ unchanged.

3. RESULTS

To evaluate the invariance properties of our feature, we generated various synthetic image sequences with known ground truths. Fig. 1 shows four subimages with additively superimposed orientations in noise, which were rotated from frame to frame by 5°. Fig. 2 shows $|\cos\beta|$ as calculated according

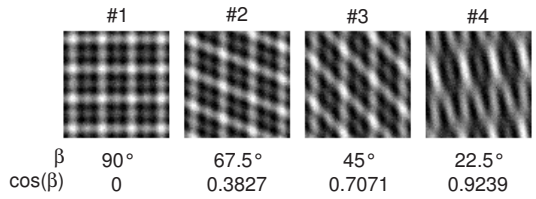

Fig. 1: First frame of a sequence showing four subimages (71×71 pixel each) with additively superimposed oriented patterns in white Gaussian noise (PSNR 28dB), where β varies from 90° to 22.5°. The structure in each subimage was rotated from one frame to the next by 5°.

to Eq. (27) vs. rotation angle for each subimage. The rotation covered in total 175°, corresponding to 35 frames. The calculation was based on a local region Ω of size 27×27 pixel placed in the image centre. For each curve, Fig. 2 also provides standard deviation and mean of our feature as estimated over all frames. Despite the presence of noise, the estimation error corresponds to less than 0.5°. Clearly, the sensitivity of our feature to rotation is very low, as expected.

Fig. 3 depicts occluding superpositions of similar patterns. Again, the feature calculation was performed within a local region of size 27×27 pixel placed in the image centre, thus always containing the occluding boundary. The plots of our feature over rotation angle are very similar to those in Fig. 2, we therefore only give standard deviations, mean values and mean difference to each true value in Tab. 1. Again, the observed sensitivity to both rotation and noise is very low.

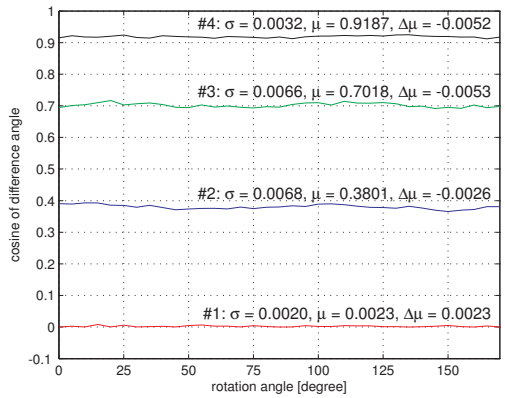

Fig. 2: Invariance feature $|\cos\beta|$ over frame number for the patterns in Fig. 1. Also given are the standard deviation σ and the mean μ, and the difference between μ and the true value of $|\cos\beta|$.

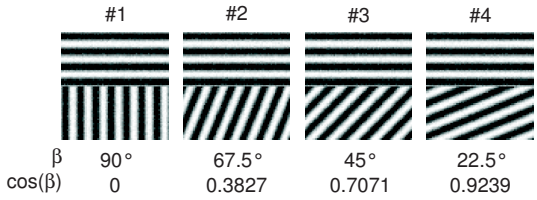

Fig. 3: Occluding patterns in noise, PSNR 28 dB.

no.	σ	μ	$\Delta\mu$
#1	0.0136	0.0154	0.0154
#2	0.0203	0.4027	0.0200
#3	0.0087	0.7197	0.0126
#4	0.0046	0.8799	-0.0440

Tab. 1: Results for the occluding patterns in Fig. 3.

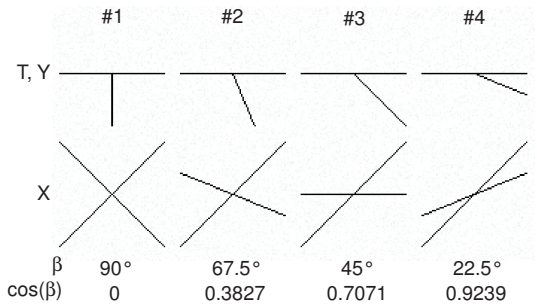

Fig. 4: T, Y and X junctions in noise, PSNR 28 dB.

Fig. 4 shows several junctions exhibiting double orientations in noise. As above, the feature calculation was carried out in a region of 27×27 pixel centred around each junction. The results of the feature values over rotation angle are given in Fig. 5, and the corresponding estimates of standard deviations, mean values and mean differences are provided in Tab. 2. The results show that such junctions pose a greater challenge than the textures above - probably because the actual orientation information covers only a minor part of the analysis region Ω. This holds in particular for the Y junction with the lowest angle β (junction #4 in the upper row of Fig. 4).

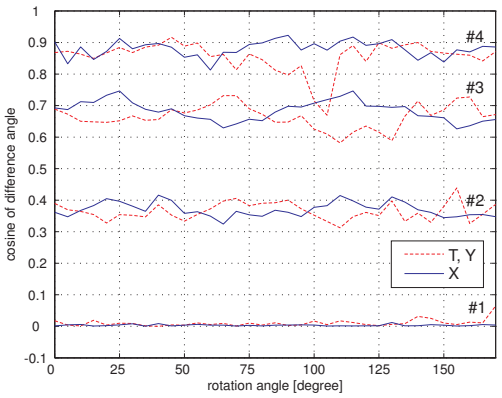

Fig. 5: Results for the rotating junctions in Fig. 4.

no.	σ	μ	$\Delta\mu$
T #1	0.0143	0.0104	0.0104
Y #2	0.0271	0.3651	-0.0176
Y #3	0.0354	0.6687	-0.0384
Y #4	0.0496	0.8549	-0.0690
X #1	0.0024	0.0030	0.0030
X #2	0.0215	0.3700	-0.0127
X #3	0.0325	0.6829	-0.0242
X #4	0.0273	0.8799	-0.0440

Tab. 2: Standard deviations and mean values computed from the curves in Fig. 5.

To evaluate the influence of the remaining fluctuations of the feature value over rotation, let us assume that each of the T and Y junctions in Fig. 4 shall be tracked by matching the features measured in a reference frame to those measured in its predecessor. The reliability of this matching process can be assessed by describing the feature fluctuations for each junction by a Gaussian distribution with mean and standard deviation as in Tab. 2. Fig. 6 shows these distributions together with the decision thresholds for optimal separation in the sense of minimum confusion error (equal a priori probabilities for the occurrence of each junction are assumed). As one would expect from Fig. 5, the overlap is largest between the distributions belonging to junctions #3 and #4. Still, the probability of an erroneous match is only 0.70%. Simply assigning each observed feature to the reference class with nearest mean would lead to an error of 0.86%.

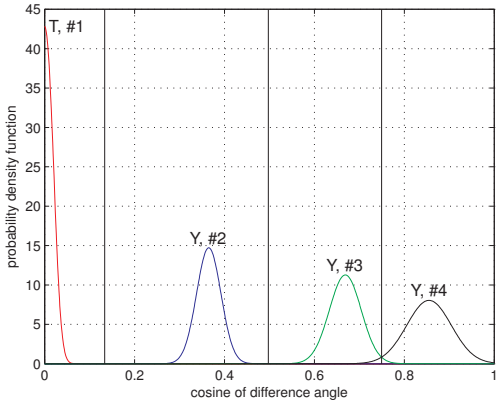

Fig. 6: Estimates of probability density functions for T and Y junctions from the parameters in Tab. 2.

4. CONCLUSION

We have derived a new invariant feature for the description of double oriented local image neighbourhoods. The derivation was partly motivated by what is known as the aperture problem in the estimation of optical flow: the matrix involved in estimating optical flow describes an important local feature, viz. orientation, but only if it is singular - in which case it cannot be used to estimate optical flow. Vice versa, if regular, the matrix permits estimation of optical flow, but it is not able to capture a local feature such as orientation any more. To characterize the underlying local image signal in this case, we have developed two superposition models assuming additively or occludingly superimposed and individually oriented subsignals. Based on these models, an extended tensor was discussed which is able to reflect the double orientation property. The eigenvector belonging the lowest eigenvalue of this extended tensor is the so-called MOP vector, which encodes the orientations of both subsignals. In bivariate images, the MOP vector has three components but, as a unit vector, it possesses only two DoF. Since invariance to rotation involves the loss of another DoF, the sought feature must be a scalar, and was intuitively found as the difference angle β between the orientations or a function of it, such as $|\cos\beta|$. We have shown how to calculate this feature directly from the MOP vector, without explicitly determining the orientations. We also extended the approach to tri- and higher-variate data, and provided a proof that indeed all scalar invariants encoded in the MOP vector are produced by β. In all cases, we have used the absolute value of $\cos\beta$ as feature, which could be directly calculated from the MOP vector components. We intentionally abstained from calculating β itself by inverting the cosine, since this involves a transcendental function the nonlinearities of which would amplify the error noise of the estimate. In addition, it could also be shown straightforwardly that the feature is invariant with respect to linear intensity scalings and the scale of the local image region.

Our experiments conducted so far with this feature confirmed the expected invariant behaviour of the feature, where we focussed on rotation of both double-oriented textured patterns and junctions. Though the results are quite promising, the development of this feature is not yet complete, and considerable research efforts remain to be carried out: first, our derivations are all based on a spatially continuous image model. In our spatially discrete implementations, we used standard filters, such as finite-difference approximations for the differentiation, which were not optimized in any way for the task at hand. The benefit of using optimized filters, such as the ones in [4, 11], remains to be investigated, as does the comparison to other rotation- and scale invariant features, such as the SIFT [15]. Another open point is the effects of sampling on the invariance with respect to scale of the neighbourhood since, as the size of local structure becomes smaller, calculation of, e.g., the derivatives will become more inaccurate. Finally, let us mention that the extension of the proposed approach to multispectral data, such as colour images, is straightforward: similarly as done for the single orientation tensor in [3], it suffices to calculate the extended tensor according to Eq. (20) on each spectral component, and to replace $\mathbf{T}$ by the sum of these.

REFERENCES

[1] T. Aach, I. Stuke, C. Mota, and E. Barth. Estimation of multiple local orientations in image signals. *Proc. IEEE ICASSP-2004*:III 553–556, Montreal, May 17–21 2004.

[2] J. Bigün and G. H. Granlund. Optimal orientation detection of linear symmetry. In *Proceedings IEEE First International Conference on Computer Vision*, pages 433–438, London, June 1987.

[3] S. Di Zenzo. A note on the gradient of a multi-image. *Computer Vision, Graphics, and Image Processing*, 33:116–125, 1986.

[4] M. Elad, P. Teo, and Y. Hel-Or. On the design of optimal filters for gradient-based motion. *International Journal of Computer Vision*, page submitted, 2002.

[5] D. A. Forsyth and J. Ponce. *Computer Vision. A Modern Approach*. Prentice-Hall, Upper Saddle River, 2003.

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

[7] M. Kass and A. Witkin. Analyzing oriented patterns. *Computer Vision, Graphics, and Image Processing*, 37:362–385, 1987.

[8] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *7th International Joint Conference on Artificial Intelligence*, pages 674–679, Vancouver, 1981.

[9] C. Mota, T. Aach, I. Stuke, and E. Barth. Estimation of multiple orientations in multi-dimensional signals. In *IEEE International Conference on Image Processing (ICIP)*, pages 2665–2668, and on CD–ROM: ISBN 0–7803–8555–1, Singapore, Oct. 24–27 2004. IEEE.

[10] C. Mota, I. Stuke, T. Aach, and E. Barth. Estimation of multiple orientations at corners and junctions. In *DAGM04: 26th Pattern Recognition Symposium*, pages 163–170, Tuebingen, Aug. 30 – Sept. 1 2004. German Association for Pattern Recognition, Springer Verlag: LNCS 3175.

[11] H. Scharr and B. Jähne. Optimization of spatio-temporal filter families for fast and accurate motion estimation. In *Image Sequence Analysis to Investigate Dynamic Processes, Lecture Notes in Computer Science*. Springer, 2003.

[12] M. Shizawa and T. Iso. Direct representation and detection of multi-scale, multi-orientation fields using local differentiation filters. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'93)*, pages 508–514, Piscataway, June 1993. IEEE.

[13] I. Stuke, T. Aach, E. Barth, and C. Mota. Analysing superimposed oriented patterns. In *6th IEEE Southwest Symposium on Image Analysis and Interpretation*, pages 133–137, Lake Tahoe, NV, March 28-30 2004. IEEE Computer Society.

[14] E. Trucco and A. Verri. *Introductory Techniques for 3D Computer Vision*. Prentice-Hall, Upper Saddle River, 1998.

[15] D. G. Lowe. Object Recognition from Local Scale-Invariant Features. In *International Conference on Computer Vision*, pages 1150–1157, Corfu, 1999. IEEE.