Single 2D Orientation Estimation via the Structure Tensor

Konstantinos G. Derpanis
kosta@cs.yorku.ca

Version 1.1

January 13, 2005

In the following note we derive the so called *structure tensor* used for the estimation of the dominant local 2D orientation. For many computer vision students the notion of a tensor is foreign. Don't be scared off!!! The *structure tensor* formulated here requires nothing more than basic knowledge of linear algebra and calculus.

Let $f(\mathbf{x})$ represent a two-dimensional image, with spatial dimensions denoted $\mathbf{x} = (x, y)$. Assume that f is oriented in a region Ω, formally,

$$f(\mathbf{x}) = f(\mathbf{x} + \mathbf{u}) \tag{1}$$

where $\mathbf{u} = (cos(\theta), sin(\theta))^\top$ describes the orientation of $f(\mathbf{x})$ parameterized by $-\frac{\pi}{2} \leq \theta < \frac{\pi}{2}$.

To find θ we pose the problem as finding the minimum gray level axis within the local neighbourhood Ω. More specifically, we find the local directional derivative that vanishes,

$$\mathbf{u} \cdot \nabla f(\mathbf{x}) = 0$$

or equivalently in matrix notation

$$\mathbf{u}^\top \nabla f(\mathbf{x}) = 0 \tag{2}$$

This can be accomplished by minimizing Eq. (2) locally in a least-squares sense, as follows,

$$E(\mathbf{u}) = \min_{||\mathbf{u}||=1} \int_\Omega (\mathbf{u}^\top \nabla f(\mathbf{x}))^2 d\Omega \tag{3}$$

Note that an additional weighting term may be added to each constraint term within the minimization for robustness purposes (e.g., a Gaussian weighting function centred in the middle of the image patch).

Expanding Eq. (3), yields,

$$
\begin{aligned}
E(\mathbf{u}) &= \min_{||\mathbf{u}||=1} \int_\Omega (\mathbf{u}^\top \nabla f(\mathbf{x}))^2 d\Omega && (4)\\
&= \min_{||\mathbf{u}||=1} \int_\Omega (\mathbf{u}^\top \nabla f(\mathbf{x}))(\mathbf{u}^\top \nabla f(\mathbf{x})) d\Omega && (5)\\
&= \min_{||\mathbf{u}||=1} \int_\Omega (\mathbf{u}^\top \nabla f(\mathbf{x}))(\mathbf{u}^\top \nabla f(\mathbf{x}))^\top d\Omega && (6)\\
&= \min_{||\mathbf{u}||=1} \int_\Omega (\mathbf{u}^\top \nabla f(\mathbf{x}))(\nabla f(\mathbf{x})^\top \mathbf{u}) d\Omega && (7)\\
&= \min_{||\mathbf{u}||=1} \int_\Omega \mathbf{u}^\top (\nabla f(\mathbf{x}) \nabla f(\mathbf{x})^\top) \mathbf{u}\, d\Omega && (8)\\
&= \min_{||\mathbf{u}||=1} \mathbf{u}^\top (\int_\Omega \nabla f(\mathbf{x}) \nabla f(\mathbf{x})^\top\, d\Omega) \mathbf{u} && (9)
\end{aligned}
$$

Let,

$$
\begin{aligned}
J &= \int_\Omega \nabla f(\mathbf{x}) \nabla f(\mathbf{x})^\top\, d\Omega && (10)\\
&= \int_\Omega \begin{bmatrix} f_x^2 & f_x f_y \\ f_x f_y & f_y^2 \end{bmatrix} d\Omega && (11)
\end{aligned}
$$

and substituting into Eq. (9),

$$
E(\mathbf{u}) = \min_{||\mathbf{u}||=1} \mathbf{u}^\top J \mathbf{u} \tag{12}
$$

Matrix J is what is usually referred to in the literature as the *structure tensor*.

One possible route to minimizing Eq. (12) is to form a Lagrange minimization. It can be shown that this minimization is equivalent to the following eigenvalue problem,

$$
J\mathbf{u} = \lambda \mathbf{u} \tag{13}
$$

where the solution we seek corresponds to the eigenvector with the smallest eigenvalue λ_2. If our model is ideally met then the rank of J is one. Violation of the single orientation model is signalled by $\lambda_1 > \lambda_2 \gg 0$. Furthermore, the eigenvalues can be used to classify regions as: *homogeneous* (i.e., no dominant orientation $\lambda_1 = \lambda_2 \approx 0$), *single orientation* ($\lambda_1 \gg 0$ and $\lambda_2 \approx 0$) and *multiple orientations* $\lambda_1 > \lambda_2 \gg 0$. Interestingly, though motivated differently the structure tensor and its subsequent eigen-analysis corresponds to the Harris corner detector analysis [1].

If one simply seeks to classify regions one can avoid the expense of explicitly calculating the eigenvalues and instead analyze the determinant K and trace H of J, formally,

$$
\begin{aligned}
K &= \det(J) = \lambda_1 \lambda_2 && (14)\\
&= (\int_\Omega f_x^2\, d\Omega)(\int_\Omega f_y^2\, d\Omega) - (\int_\Omega f_x f_y\, d\Omega) && (15)\\
H &= \mathrm{trace}(J) = \lambda_1 + \lambda_2 && (16)\\
&= (\int_\Omega f_x^2\, d\Omega) + (\int_\Omega f_y^2\, d\Omega) && (17)
\end{aligned}
$$

where in the ideal case,

- **Homogeneous**: $H = 0$ (ideal case $\lambda_1 = \lambda_2 = 0$)
- **Single Orientation**: $H > 0$ and $K = 0$
- **Multiple Orientations**: $H > 0$ and $K > 0$ (condition used for Harris Corner Detector [1])

This concludes the formulation of the structure tensor for the 2D dominant orientation case. Importantly, this basic formulation can be extended to various other multidimensional problems. For example, the problem of optical flow estimation can be formulated as a tensor of the spatiotemporal image structure [2]; it turns out that the structure tensor for optical flow yields an estimate equivalent to the total least squares estimate [3]. In fact, any linear partial differential equation can be written as a structure tensor [4]. For example, in [4] the authors demonstrate an "extended" structure tensor for motion analysis that can handle brightness changes other that those caused by movement in the scene (i.e., violations of the brightness constancy assumption).

References

[1] C. Harris and M.J. Stephens. A combined corner and edge detector. In *Alvey Vision Conference*, pages 147–152, 1988.

[2] H. Haußecker and H. Spies. *Handbook of Computer Vision and Applications*, volume 2, chapter 13. Academic Press, 1999.

[3] H. Haußecker, H. Spies, and B. Jähne. Tensor-based image sequence processing techniques for the study of dynamical processes. In *International Symposium On Real-time Imaging and Dynamic Analysis*, pages 704–711, 1998.

[4] H. Spies and B. Jähne. A general framework for image sequence analysis. In *Fachtagung Informationstechnik*, pages 125–132, 2001.